CUK-H06047-27-P4783

# THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

27

FEBRUARY 1988

VOL. 58, PART 1

# THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by
Professor HAZEL FRANCIS

Assistant Editors:
Dr. H. C. M. CARROLL
Prof. C. W. DESFORGES
Dr. D. FONTANA
Dr. F. J. TAYLOR

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Prof. C. W. DESFORGES

**Advised by the Editorial Committee:**
Professor G. BROWN
Ms. I. LUNT
Dr. C. LEACH
Professor P. W. ROBINSON
Mr. E. WILKINSON

P 4783

Vol. 58, Part 1

*The British Journal of Educational Psychology* is issued by the British Psychological Society once a term, in February, June and November. The subscription is £20.00 a year to the UK, £22.00 overseas, or £7.00 each number.

*Subscriptions and Orders* should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

*Papers for publication* in duplicate should be sent to Professor Hazel Francis, University of London Institute of Education, 24-27 Woburn Square, London, WC1H 0AA. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to: Prof. C. W. Desforges, University of Exeter, St. Luke's, Heavitree Road, Exeter, EX1 2LU.

*Advertisements* for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

*All rights reserved*

© 1988. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY LINDSAY & CO LTD , EDINBURGH

370.150
B777

# P. E. VERNON

## Teacher, Guru and Friend

Philip Vernon who died in Calgary, Alberta, on 28th July, 1987, was a remarkable man, whose contribution to our understanding of educational psychology was immeasurable. It is a great honour to me, as one of his many former students, to be asked to write this appreciation.

One of the pioneers in mental testing, attainment and personality, he contributed widely to the nature-nurture debate. While some of his very considerable output might seem to be of less direct relevance today when attitudes to intelligence testing have rightly become more sceptical, the Vernon approach which characterised his life-work is an exemplar to all in terms of its rigour, balance and above all in its humane and sympathetic approach.

It is easy, particularly in this present age, to point to the limitations of the psychometric approach, but Vernon was always the first to recognise these, seeking to improve techniques and, above all, ready to look for factors beyond his current theses to extend his knowledge and understanding of the subject.

His output was characterised by painstaking scholarship. One has only to examine the bibliographies accompanying his major survey works to appreciate the scale of this. Moreover, the judgments formed were never precipitate but were balanced and the product of a scrupulous and open-minded evaluation of all the evidence available.

His reputation owes much to the books he produced in a 20-year period as Professor of Educational Psychology at the London Institute of Education: *The Structure of Human Abilities* (1950), *Personality Tests and Assessments* (1953), *Intelligence and Attainment Tests* (1960), and *Personality Assessment* (1963). Comprehensive and balanced surveys of the then available materials and evidence, these became standard works of reference for several generations of students, teachers and researchers, both in the UK and overseas.

During his years at the Institute particularly, Philip Vernon travelled widely attending conferences, pursuing research or on lecture tours. Indeed, he reckoned once that he had taught in well over 25 countries world-wide. This experience made him increasingly sensitive to the influence of cultural factors on intelligence. That interest was clearly reflected in *Intelligence and the Cultural Environment* (1969), which he regarded as one of his most important studies, and in *Intelligence, Heredity and the Environment* (1979).

At an age when many might have been content to prepare for retirement, Philip Vernon resigned his chair at the Institute in 1968 to move to Canada to take up a similar post in Calgary, Alberta. Even eventual retirement in 1975 did not cause his activity to diminish, and he continued as Professor Emeritus to supervise research students, and continued his scholarship to the very end. His last book, *The Abilities and Attainments of Orientals*, was published in 1985, in his eightieth year.

For those tempted to ask what Philip Vernon's legacy is, then for me the answer is an easy and unequivocal one. It lies in his approach. He was ever flexible of mind and balanced in his views. He followed research and debate assiduously, reading and listening widely. His was an enquiring mind, an open one, that went relentlessly in pursuit of increased understanding of the world around him. Those values were ones Philip Vernon instilled not only through his writings but also in many generations of grateful educationists across the world who sat at his knee at one time or another.

GAJENDRA K. VERMA,
*University of Manchester.*

P4783

*Br. J. educ. Psychol.*, 58, 1-14, 1988

# ENHANCING AND UNDERMINING INTRINSIC MOTIVATION: THE EFFECTS OF TASK-INVOLVING AND EGO-INVOLVING EVALUATION ON INTEREST AND PERFORMANCE

BY RUTH BUTLER
*(Hebrew University of Jerusalem)*

SUMMARY. This study was designed to test the hypothesis that intrinsic motivation will be differentially affected by task-involving and ego-involving evaluation, and that provision of both kinds of evaluation will promote ego-involvement rather than task-involvement. Twelve classes of fifth and sixth grade pupils were randomly assigned to one of three feedback conditions. Pupils received either ego-involving numerical grades or task-involving individual comments or both after performing interesting tasks, one convergent and one divergent, on each of two sessions. Interest and performance for 132 randomly selected pupils of high or low school achievement were measured at pre-test, during the manipulation and at a third session, when no further evaluation was anticipated. As hypothesised, interest and performance on both tasks at both levels of school achievement were highest after comments, both when further comments were anticipated and when they were not. Grades and grades plus comments had similar and generally undermining effects on both interest and performance, although high achievers who received grades maintained high interest and convergent thinking when further grades were anticipated. These results are discussed in terms of the contribution of this distinction between task and ego-involvement to further understanding of intrinsically motivated activity.

## INTRODUCTION

RECENT research on intrinsic motivation has consistently found that rewards undermine subsequent interest for initially attractive tasks, apparently by promoting attributions of task engagement to the reward rather than to pleasure in the activity itself (see recent reviews by Lepper, 1983; Morgan, 1984). However, less attention has been directed to clarifying how initial interest can be maintained, or possibly even enhanced. Cognitive evaluation theory (Deci, 1975; Deci and Ryan, 1980) maintains that interest will not be undermined by rewards which are perceived as providing positive information about competence rather than as sources of control. However, Harackiewicz, *et al.* (1984) note that this prediction has not received consistent empirical support and suggest that this may be because the mechanism by which positive information is supposed to enhance interest remains unclear.

Similar concerns have been expressed by several researchers who argue that theories which focus primarily on the implications of exogenous perceptions of causality for task resumption have shed little light on the ways in which the experience and process of task engagement may differ under intrinsic or extrinsic conditions (Csziksentmihalyi, 1975; Maehr, 1976, Condry, 1977; deCharms, 1983). Thus it is also not clear what relationship, if any, should exist between interest, or intrinsic motivation, and performance. While early studies (Kruglanski *et al.*, 1971; Greene and Lepper, 1974) suggested that rewards may undermine creative or divergent thinking, these findings seemed compatible with general behaviour theory and had little impact on conceptualisations of intrinsic motivation. Moreover, later studies tended to find that performance was largely unrelated either to the presence or absence of rewards or to increases or decreases in subsequent interest (see review by Morgan, 1984).

A somewhat clearer picture of intrinsically motivated activity seems to emerge from the approaches of deCharms (1968) and Nicholls (1983, 1984). Both authors distinguish between extrinsic, task-involved and ego-involved motivational orientations according to their characteristic foci of attention — on external goals, mastery and self-worth respectively. Intrinsic, or task-involved, motivation is characterised by the concern to improve mastery vis-à-vis task demands and/or prior performance and should be maintained as long as the task is perceived as relevant to the ongoing development and assessment of individual mastery (Nicholls, 1983, 1984). Bandura too emphasises the role of information seeking about competence in maintaining self-regulated motivation (Bandura and Schunk, 1981; Bandura, 1982).

Such positive definitions of intrinsic motivation in terms of competence development and assessment seem to have clearer implications for interest enhancement and for the relation between interest and performance than the largely negative definition of those analyses which emphasise the absence of extrinsic constraints. Thus one can predict that continuing task-involvement and interest should importantly depend on the availability of information about performance, since most activities in both experimental and applied settings do not in themselves provide a basis for the self-evaluation of competence (Bandura and Cervone, 1983). In addition, the focus of attention on mastery suggests that task-involvement will also enhance performance. Moreover, interest and performance should be maintained, at least in the short term, even when no further information is anticipated, since its provision in the past will have enhanced perceptions of the task as relevant to developing mastery and will have promoted the formation of internal standards for evaluating performance and setting goals.

This analysis seems to have clear implications for educational settings, which continually evaluate student performance. However, not all information should enhance interest, even when the task in question is attractive. Controlling information — if you engage in this activity you will be rewarded — should result in a shift from the initial task-involved orientation to the extrinsic one studied by means-end analyses. Normative information, such as that employed as an "intrinsic" motivational condition in studies guided by cognitive evaluation theory, should also promote a shift from initial task-involvement, in this case to a self-worth or ego-involved motivational orientation. Here attention is focused on demonstrating high ability or masking low ability and there is self-esteem-based pressure to achieve positive and avoid negative outcomes (Nicholls, 1984). While ego-involvement has been studied mainly in the context of achievement behaviour, rather than of continuing interest, research findings indicate that task choice, satisfaction and performance are importantly mediated by perceived ability in ego-involving settings (see review by Nicholls, 1984).

This distinction between task-involved and ego-involved motivation suggests several predictions as to the effects of different kinds of information on both performance and interest. Individualised information which relates specifically to aspects of the task which have or have not been mastered should maintain initial task-involvement in both more and less successful students. Thus one can predict high interest and enhanced performance, even when no further information is anticipated. In contrast, the anticipation of ego-involving information, such as the normative grades prevalent in schools, should have differential effects at high and low levels of perceived ability. While grades may maintain initial levels of interest and performance in high achievers who expect these to be self-enhancing, these should be undermined in low achievers. However, the prediction that task and ego-involvement will yield similar levels of performance in high achievers who anticipate further evaluation may be valid only for tasks requiring convergent thinking. Thus findings that divergent or creative thinking is undermined by ego-involving cues such as time limits (Wallach and Kogan, 1965) and the anticipation of evaluation (Amabile, 1979) suggest that even able subjects will perform better on such tasks when they expect comments than when they expect grades. Finally, one can predict that both interest and performance will decline even among high achievers if no further normative information is anticipated, since the task will no longer be perceived as relevant to demonstrating high ability.

This may be the psychological mechanism behind observations that pupils often seem more interested in their own and their friends' grades than in learning. Nicholls' findings (1978) on developmental transitions in the concept of ability suggest that normative evaluation may be less ego-involving in the early grades, when pupils do not perceive ability as a stable trait and tend to judge mastery relative to task demands and past performance rather than to others' outcomes. However, by fourth grade most pupils have achieved a differentiated concept of ability as a stable trait best assessed by comparison with others. At this point grades should induce a shift from task- to ego-involvement. If so, one can predict further that, just as combining initial interest with extrinsic incentives results in an extrinsic orientation, combining task and ego-involving information should induce ego-involvement, at least after fourth grade. While many teachers seem to feel that any negative effects of grades can be overcome by adding a personal comment, the present analysis suggests that this practice should affect interest and performance much as do grades alone.

Verification of these predictions as to the motivational impact of individual comments, normative grades and comments plus grades necessitated a complex design which would compare their effects on interest and performance before, during and after receipt of information, for both a convergent and a divergent task and for both high and low levels of perceived ability. In the present study, fifth and sixth grade pupils, randomly sampled from either high or low levels of school achievement, received two interesting tasks, one convergent and one divergent, on each of three sessions. No evaluation was anticipated on Session 1 (pre-test); at Session 2, pre-test tasks were returned with appropriate comments, numerical grades, or both and subjects expected to receive similar evaluation on new tasks; at Session 3 (post-test) evaluation was received for Session 2 tasks, but not anticipated for present ones. Interest and performance were measured at each session.

The experimental hypotheses can be summarised as follows:

(a) Post-test interest and performance on both tasks will be highest after receipt of comments at both levels of school achievement.

(b) High achievers will score similarly on Session 2 interest and convergent thinking in all groups, while low achievers will score highest after comments; both high and low achievers will score highest on immediate divergent thinking after comments.

(c) Subjects who received comments alone will recall these better than subjects who also receive a grade; changes in performance from pre-test to post-test will be related to the content of the comments received earlier in the Comments but not in the Grades + Comments condition.

(d) Patterns of interest and performance on Sessions 2 and 3 will be similar in the Grades and Grades + Comments conditions.

## METHOD

### *Sample*

The sample comprised 132 fifth and sixth grade Jewish Israeli pupils (68 boys, 64 girls, mean age 11·10) from 12 randomly selected classes (out of 23) in four elementary city schools serving predominantly middle-class populations. Four classes were randomly assigned to each of the three experimental conditions. Although all pupils participated in the experiment, data were analysed only for 44 pupils in each condition, randomly selected from all those whose average grade in language and mathematics in their most recent report card was in the top or bottom 25 per cent for their class. Thus there were 22 high achievers and 22 low achievers in each experimental condition.

### *Instruments*

These consisted of three work booklets containing the experimental tasks for Sessions 1, 2 and 3 respectively. Each booklet contained two tasks, A and B. For Sessions 1 and 3 Task A consisted of a design of five concentric circles, with three or four letters of the Hebrew alphabet printed in each band. Subjects were asked to construct as many words as they could from these letters, starting with a letter in the centre and then adding a letter from one or more successive bands. Task B consisted of two examples from the divergent thinking "uses" test (Torrance and Templeton, 1963). Tasks for Session 2 were slightly different, so as to reduce boredom and practice effects. For Task A, children were asked to construct words from the letters of a long word. Task B consisted of the divergent thinking "circles" test. A pilot study established that pupils of this age found the tasks interesting and that the tasks given on Sessions 1 and 3 yielded equivalent levels of performance.

In addition, an interest questionnaire given after each session asked pupils to rate their interest and enjoyment of the tasks and the degree to which other pupils would find them interesting on seven-point scales. After Sessions 1 and 3 subjects were also asked to state how many additional tasks they would like to receive (from 1 to 7) and after Session 3 also to recall the evaluation received on Session 2.

### *Feedback conditions*

Feedback was given after Sessions 1 and 2 as follows. Unless clearly inappropriate, each child was given similar feedback after each session to reduce the possibility that discrepancies would in themselves affect post-test measures.

(i) *Comments group.* Feedback consisted of one sentence, which related specifically to the performance of the individual child but did not include any information beyond that specified in the criteria for success given to all subjects in the general instructions before each session. For each session, comments took one of the following standardised forms, with slight variations in wording but not content:

> *Task A:* You thought of quite a few correct words; maybe it is possible to think of more short words/more long words/even more words.
>
> *Task B:* You thought of quite a few interesting uses (or "ideas" for the circles task); maybe it is possible to think of more uses (or ideas. . .)/more different uses/more unusual uses which other children may not think of.

Thus the first phrase for each task was the same, but was followed by one of three different suggestions for improvement.

(ii) *Grades group.* Final performance scores for all pupils, including those of average ability, were computed as described below. These scores were then converted into grades so as to follow a normal distribution ranging from 40 to 99.

Thus high achievers tended to receive relatively high grades and low achievers relatively low ones.

(iii) *Grades + Comments group.* Subjects were given both a grade and a comment, as described above.

### *Procedure*

The experiment consisted of three sessions. Session 1 was conducted in one day and Sessions 2 and 3 two days later, with an interval of three hours between them. Sessions were conducted during regular school hours by one of two undergraduate students in psychology. Session 1 instructions, identical for each feedback group, explained that the experimenters had constructed some tasks and were interested in seeing how different children answered them; they hoped that the children would enjoy doing them. Instructions for Task A were then read out. These included rules regarding acceptable words and criteria for successful performance (according to the correctness, number and length of words). Pupils were asked to begin and after 10 minutes to stop and turn to Task B. Instructions for the "uses" task, adapted from Torrance and Templeton (1963), also included criteria for success (according to the number, variety and originality of responses). After 10 minutes the tasks were collected and the interest questionnaire distributed.

In Session 2, two days later, these tasks were returned. Pupils in Group 1 were told that each had been given an appropriate comment on his/her performance. Pupils in Group 2 were told that each had been given a grade which indicated how they had done relative to other pupils in the class, and those in Group 3 that each had been given a grade and a comment. All pupils were instructed to look at the tasks, to see how they had done, after which they would be given new tasks. Pupils in Group 1 were told that they would receive a comment on the new tasks too, in Group 2 that they would receive a grade and in Group 3 that they would again receive a grade and a comment. Booklet 2 was then distributed and the procedure for Tasks A and B was followed as for Session 1. The interest questionnaire was then given out. In Session 3, three hours later, booklet 2 was returned with the evaluation appropriate to each group and booklet 3 was distributed. Pupils were given a few minutes to look through booklet 2 and were then told that the experimenters would like them to try out some new tasks. They were told that these would not be evaluated or returned. The procedure for Tasks A and B and for the interest questionnaire was identical to previous sessions. On completion, pupils were engaged in a discussion about the experiment.

### *Scoring*

*Task A.* Note was made of the number of short (two or three letter words) and long (four or more letter) words and of the total number of words. A final score was computed by awarding each short word one point and each long one two points.

*Task B.* Scoring was according to the categories defined by Torrance and Templeton (1963). Counts were made of the number of responses (fluency), categories (flexibility), elaborated responses and original responses, where originality was defined as a non-bizarre idea which appeared in no more than 10 per cent of the protocols. A final score was computed from the sum of scores in each category, which for the "uses" task were averaged over the two examples given in Sessions 1 and 3. Originality was given a weight of 2 and the other components a weight of 1. Tasks were scored by two judges working independently. Inter-judge agreement for final scores was high: $r = 0{\cdot}91$ for Session 1, $r = 0{\cdot}89$ for Session 2 and $r = 0{\cdot}93$ for Session 3.

## RESULTS

### *Performance measures*

Table 1 presents mean scores by feedback condition and achievement level for final Task A and Task B performance scores at Sessions 1, 2 and 3. A two-way analysis of variance by feedback condition and school achievement yielded no significant main or interaction effects for feedback condition on Session 1 final scores. The significant main effects for school achievement, F(1, 125) = 13·02, $P < 0·001$ for Task A and F(1, 125) = 6·01, $P < 0·05$ for Task B, confirmed that pupils whose school grades were high performed better on both tasks than pupils whose grades were low. Since preliminary analyses yielded no significant main or interaction effects for sex on any of the dependent variables, data for the sexes were combined for further analysis. Final Task A and Task B scores on Sessions 2 and 3 were then analysed with 3 × 2 (Feedback Condition × School Achievement) ANOVAS. When appropriate, specific hypotheses were then tested using orthogonal planned contrasts.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR FINAL SCORES AT EACH SESSION BY FEEDBACK CONDITION AND SCHOOL ACHIEVEMENT

| | | Comments group | | Grades group | | Grades plus Comments group | |
|---|---|---|---|---|---|---|---|
| | | High | Low | High | Low | High | Low |
| *Task A* | | | | | | | |
| Session 1 | M | 18·77 | 10·14 | 19·64 | 9·86 | 19·60 | 9·64 |
| | SD | 6·63 | 5·02 | 8·67 | 4·88 | 5·80 | 5·64 |
| Session 2 | M | 25·36 | 17·86 | 24·95 | 12·50 | 16·77 | 6·55 |
| | SD | 9·33 | 8·70 | 9·91 | 7·86 | 5·61 | 4·73 |
| Session 3 | M | 24·27 | 13·50 | 16·45 | 8·59 | 11·82 | 5·82 |
| | SD | 6·70 | 7·79 | 9·34 | 6·41 | 5·26 | 5·40 |
| *Task B* | | | | | | | |
| Session 1 | M | 19·36 | 10·27 | 19·68 | 10·32 | 18·68 | 10·12 |
| | SD | 5·35 | 4·01 | 7·36 | 3·71 | 6·60 | 3·63 |
| Session 2 | M | 25·59 | 17·64 | 16·09 | 12·59 | 14·36 | 11·95 |
| | SD | 9·39 | 10·16 | 6·63 | 5·94 | 5·17 | 3·71 |
| Session 3 | M | 24·95 | 14·27 | 14·91 | 8·50 | 14·95 | 9·18 |
| | SD | 5·38 | 4·61 | 6·63 | 3·07 | 5·83 | 3·40 |

### *Session 2*

*Task A*. It was hypothesised that while high achievers would score higher than low achievers in all conditions, low achievers would score highest after comments and the scores of high achievers would be similar in all feedback conditions. Although the interaction effect was not significant, a planned comparison for low achievers who received comments versus those who received grades or grades plus comments was significant, F(1, 125) = 14·54, $P < 0·001$; however, so too was that comparing the scores of low achievers after grades and grades plus comments, F(1, 125) = 5·84, $P < 0·05$. While the former comparison was not significant for high achievers, that comparing scores in the grades and grades plus comments conditions was, F(1, 125) = 10·85, $P < 0·001$. Thus both high and low achievers scored higher in the grades than in the grades plus comments condition.

*Task B.* As predicted, the results yielded a significant main effect for feedback condition, F(2, 125) = 17·67, P < 0·001 and a non-significant interaction effect. Planned contrasts confirmed that over both levels of achievement pupils who received comments scored highest on this divergent thinking task, F(1, 125) = 33·96, P < 0·001, while there was no significant difference in the scores of those who received grades or grades plus comments. Table 1 indicates that these trends were found at both levels of achievement. It is interesting that although the main effect for school achievement was significant, F(1, 125) = 14·15, P < 0·001, low achievers who received comments scored higher than high achievers in the other conditions (see Table 1).

### *Session 3*

It was hypothesised that scores for both tasks would be highest after comments at both levels of achievement and that there would be no significant differences in scores after grades and grades plus comments.

*Task A.* The ANOVA yielded the predicted main effects for feedback condition F(2, 125) = 29·60, P < 0·001 and school achievement, F(1, 125) = 57·48, P < 0·001 and a non-significant interaction effect. The appropriate contrasts confirmed that scores over both levels of achievement after comments were higher than combined scores after grades or grades plus comments F(1, 125) = 98·68, P < 0·001. However, the comparison between scores in the latter groups was also significant, F(1, 125) = 11·8, P < 0·001. Thus, as in Session 2, pupils who received grades scored higher than pupils who received both a grade and a comment (See Table 1).

*Task B.* The ANOVA yielded the predicted main effects for feedback condition, F(2, 125) = 44·36, P < 0·001 and school achievement, F(1, 125) = 75·73, P < 0·001. The appropriate contrasts confirmed that divergent thinking scores over both levels of achievement were highest after comments, F(1, 125) = 160·1, P < 0·001 and did not differ significantly after grades and grades plus comments. As in Session 2, low achievers who received comments scored higher than high achievers in the other conditions (see Table 1).

### *Interest measures*

Preliminary two-way analyses of variance yielded no significant main or interaction effects for Session 1 ratings for any of the four interest questions. Since intercorrelations between ratings for the three measures which tapped perceived interest were high: r (132) = 0·68 to 0·79 for Session 1; 0·72 to 0·82 for Session 2 and 0·71 to 0·81 for Session 3, responses to these questions were averaged into a single measure. The question asking how many more tasks pupils would like to receive, which was asked only after Sessions 1 and 3, was analysed separately because of its different face content. Table 2 presents the between-session correlations for these measures within each feedback condition. As hypothesised, these were considerably higher in the Comments than in the Grades or Grades plus Comments conditions. Mean interest ratings for Sessions 2 and 3, presented as Table 3, were then analysed using two-way ANCOVAS, with the corresponding Session 1 ratings as the covariate.

### *Session 2*

As expected, the analysis for the composite measure of perceived interest yielded a significant effect for the interaction of feedback condition with school achievement, F(2, 125) = 27·44, P < 0·001. The significant effect for the covariate, F(1, 124) = 37·21, P < 0·001 indicated that interest ratings on Sessions 1 and 2 were linearly related. Orthogonal planned contrasts performed on the adjusted cell means confirmed that the combined interest of high achievers who received grades or grades plus comments was higher than that of low achievers in these conditions, $F(1, 125) = 13{\cdot}56$, $P < 0{\cdot}001$, while the interest of high and low achievers in the comments group did not differ significantly. Low achievers expressed most interest after comments, while high achievers expressed similar interest in all feedback conditions (see Table 3).

TABLE 2

INTERCORRELATIONS BETWEEN INTEREST MEASURES FOR EACH EXPERIMENTAL SESSION, BY FEEDBACK CONDITION

| | Feedback Condition | | |
|---|---|---|---|
| | Comments group | Grades group | Grades plus Comments group |
| Perceived Interest | | | |
| Session 1 with Session 2 | 0·31** | 0·18 | 0·16 |
| Session 1 with Session 3 | 0·42** | 0·19 | 0·21* |
| Session 2 with Session 3 | 0·36** | 0·24* | 0·19 |
| Extra Tasks Requested | | | |
| Session 1 with Session 3 | 0·43** | 0·12 | 0·11 |
| N | 44 | 44 | 44 |

*P < 0·05 **P < 0·001

TABLE 3

MEANS AND STANDARD DEVIATIONS FOR INTEREST RATINGS AT EACH SESSION BY FEEDBACK CONDITION AND SCHOOL ACHIEVEMENT

| | | Comments group | | Grades group | | Grades plus Comments group | |
|---|---|---|---|---|---|---|---|
| | | High | Low | High | Low | High | Low |
| Perceived Interest Composite | | | | | | | |
| Session 1 | M | 6·00 | 6·05 | 6·35 | 6·27 | 6·42 | 6·27 |
| | SD | 0·90 | 0·95 | 0·88 | 0·83 | 0·66 | 0·55 |
| Session 2 | M | 6·33 | 6·23 | 5·62 | 3·52 | 5·71 | 4·03 |
| | SD | 0·46 | 0·65 | 0·56 | 1·02 | 0·49 | 0·97 |
| | $M^a$ | 6·43 | 6·31 | 5·57 | 3·50 | 5·63 | 4·01 |
| Session 3 | M | 6·61 | 6·45 | 4·47 | 4·39 | 4·95 | 4·85 |
| | SD | 0·42 | 0·62 | 1·50 | 1·15 | 1·36 | 1·18 |
| | $M^a$ | 6·71 | 6·52 | 4·42 | 4·37 | 4·88 | 4·93 |
| Extra Tasks Requested | | | | | | | |
| Session 1 | M | 5·80 | 5·64 | 6·13 | 5·84 | 6·18 | 5·91 |
| | SD | 1·22 | 1·55 | 0·90 | 1·11 | 1·01 | 1·02 |
| Session 3 | M | 6·48 | 6·02 | 3·50 | 3·32 | 4·09 | 4·00 |
| | SD | 0·71 | 1·29 | 1·53 | 1·38 | 2·02 | 1·35 |
| | $M^a$ | 6·49 | 6·24 | 3·37 | 3·31 | 3·94 | 3·96 |

$M^a$ — M is adjusted for Session 1 covariant.

*Session 3*

As hypothesised, the analysis for perceived interest yielded a significant effect only for feedback condition, $F(2, 125) = 49{\cdot}54$, $P < 0{\cdot}001$. Planned contrasts performed on the adjusted cell means confirmed that the combined interest of high and low achievers was highest after comments, $F(1, 125) = 147{\cdot}6$, $P < 0{\cdot}001$ and that the comparison between the combined ratings of high and low achievers in the grades and grades plus comments conditions was not significant. The analysis for the number of extra tasks requested also yielded only a significant main effect, $F(2, 125) = 66{\cdot}66$, $P < 0{\cdot}001$. Again, the comparison over both levels of achievement between the comments and the other conditions was significant, $F(1, 125) = 245{\cdot}4$, $P < 0{\cdot}001$, while that between the grades and grades plus comments condition was not.

*Relation between Session 1 comments and Session 3 performance*

The theoretical framework developed above implied that performance on Session 3 would be more closely related to the content of the comments received after Session 1 in the Comments than in the Grades plus Comments condition. Pupils in the Comments and Grades plus Comments conditions were further divided into three groups according to which of the three possible comments they had received after each task on Session 1. Table 4 presents the percentages of subjects whose Session 3 scores on the various components of each task were higher than at Session 1.

TABLE 4

PERCENTAGES OF PUPILS WHO IMPROVED ON EACH TASK COMPONENT BY COMMENT RECEIVED ON SESSION 1

Suggestion for Improvement

*Task A*

| | Short words | | Long words | |
|---|---|---|---|---|
| Positive change for Session 3 — Session 1 | Comment | Grades plus Comments | Comment | Grades plus Comments |
| Short words | 87 | 31 | 42 | 24 |
| Long words | 27 | 25 | 95 | 35 |

*Task B*

| | More ideas | | More different ideas | | More unusual ideas | |
|---|---|---|---|---|---|---|
| Positive change for Session 3 — Session 1 | Comment | Grades plus Comments | Comment | Grades plus Comments | Comment | Grades plus Comments |
| Fluency | 100 | 22 | 46 | 31 | 46 | 10 |
| Flexibility | 50 | 22 | 92 | 31 | 38 | 30 |
| Originality | 22 | 22 | 8 | 13 | 77 | 20 |

In the Comments condition one finds a clear relation between the content of the Session 1 comment and improvement on specific aspects of each task. Thus 87 per cent of the pupils told after Session 1 that they had written several correct words but could have written more short words did write more short words on Session 3, while only 27 per cent wrote more long words. Similarly, 92 per cent of subjects told that they could have written more long words did so, as compared with 42 per cent who wrote more short ones. As Table 4 indicates, the same pattern of differential improvement by comment received was found also for Task B. While these results are hardly surprising in themselves, they are of considerable interest when compared with those for pupils who received grades plus comments. As could be expected from the data for final scores reported above, overall fewer pupils in this condition improved their scores on specific components of task performance. However, it is striking that similar percentages of subjects improved their scores on each component, regardless of which specific suggestion for improvement they had received on Session 1. Thus, for example, 35 per cent of the subjects who were told that they could have written more long words did so, but this proportion is similar to the 25 per cent who had been told after Session 1 that they could have written more short words. Table 4 reveals the same pattern of non-differential improvement also for Task B.

*Recall of evaluation*

This was tapped by asking pupils to recall the evaluation they had received for Task B on Session 2. Task B was chosen since the comments for Session 2 were worded somewhat differently than those for Session 1. It was emphasised that we were interested in seeing how well subjects could remember their evaluation in order to reduce the possibility that pupils might recall their evaluation but refrain from writing this down. The number of pupils who recalled their grade and/or comment accurately in each feedback condition are presented as Table 5. As hypothesised, almost all pupils who received a grade or both a grade and a comment remembered the grade. Somewhat fewer pupils who received only a comment remembered all of it, but all but two recalled at least one component. However, only 45 per cent of the pupils who received both a grade and a comment remembered even part of the comment, although they were specifically asked to recall both the grade and comment.

TABLE 5

RECALL OF TASK B SESSION 2 EVALUATION IN EACH FEEDBACK CONDITION

| | Comments group | Grades group | Grades plus Comments group |
|---|---|---|---|
| Percentage of subjects recalling evaluation | | | |
| Comment alone | 84 (95[a]) | — | 5 (11[a]) |
| Grade alone | — | 89 | 53 |
| Grade plus Comment | — | — | 23 (34[a]) |

[a] — per cent who recalled only one component of the comment.

## DISCUSSION

The results confirm the importance of distinguishing between task-involvement and ego-involvement when investigating intrinsic task motivation. As hypothesised, both high and low achievers who received comments continued to express high interest both on Session 2, when they anticipated further comments, and at post-test, when they did not. Similar patterns were received also for their performance on both convergent and divergent tasks. Additional findings that interest was slightly higher at post- than at pre-test, that performance on both tasks was considerably higher at post-test and that later interest was highly correlated with initial interest provided further support for the prediction that initial task-involvement would be maintained in this condition.

In contrast, it was hypothesised that normative grades would cause a shift from the initial task-involved orientation at pre-test to an ego-involved orientation in later sessions, not only when grades alone were provided, but also when these were given in conjunction with task-involving comments. As hypothesised, in both conditions initial interest was less predictive of interest on both later sessions than after comments. In addition, the results confirmed the predictions that immediate interest and convergent thinking would be maintained for high achievers and undermined for low achievers, and that both immediate divergent thinking and subsequent interest and performance on both tasks would be undermined at both levels of achievement. Thus task-involving feedback does seem to have different effects on both interest and performance than ego-involving feedback. In addition, while many teachers seem to feel that any negative effects of grades can be ameliorated by adding a personal comment, the above results suggest that this practice too will induce an ego-involved orientation. This suggestion was further supported by the finding that pupils did indeed tend to recall the grade rather than the comment.

While the differential effects on interest of comments as compared with grades and grades plus comments are consistent with the above characterisation of these conditions as task and ego-involving respectively, it is important to consider whether they could not have been predicted also by cognitive evaluation theory. Both anecdotal evidence (Holt, 1964) and some research findings (Maehr and Stallings, 1972; Harter, 1978) suggest that grades are perceived as potent sources of control over learning. The present findings of reduced subsequent interest after grades as compared with comments is consistent with cognitive evaluation theory's prediction that controlling information will undermine interest relative to non-controlling information, and Butler and Nisan (1986) did indeed interpret their findings that grades yielded lower subsequent interest than task-related comments within this framework. Moreover, if grades are perceived primarily as salient extrinsic incentives, it is not surprising that they continued to undermine subsequent interest in the present study also when they were given in conjunction with comments.

However, in this case, the two grades conditions should also have undermined interest relative to pre-test on Session 2 and not only on Session 3, since at both points the grades should have provided sufficient justification for task engagement. Moreover, these decrements should have been more marked for low achievers who received poor grades (negative information) than for high achievers who received high ones, not only on Session 2 but also on Session 3. Finally, the comments provided, while probably less controlling than grades, seem at least as controlling as the phrase: "Good. You should keep up the good work" which Ryan (1982) found to undermine interest. In addition, they provided mixed rather than clearly positive information. Thus conceptualising these feedback conditions as task-involving and ego-involving respectively seems to provide a more comprehensive framework than the controlling-informational distinction of cognitive evaluation theory for interpreting both the differential effects of grades on immediate as compared with subsequent interest and the consistently high interest expressed after comments on all sessions. Furthermore, while normative information about outcomes may undermine subsequent interest less than tangible rewards, the implication that this does indeed seem to promote ego-involvement rather than task-involvement reinforces the reservations noted above regarding its conceptualisation as an "intrinsic" motivational condition in many studies.

It seems less relevant to ask how cognitive evaluation theory would interpret the findings for performance, since this has been concerned primarily with predicting subsequent interest. Those studies which have measured performance have tended to compare performance during the manipulation with interest at post-test. These studies have yielded inconsistent results (Morgan, 1984), which have led some researchers (cf. Deci and Ryan, 1980) to the somewhat counter-intuitive conclusion that there is no necessary relation between performance and interest. In contrast, our results suggest that different kinds of feedback, conceptualised as task and ego-involving respectively, do indeed affect both interest and performance in similar ways. However, this becomes clear only when these are both assessed at the same points in the motivational process.

In addition, the present findings also provide some insights as to the ways in which task-involved and ego-involved motivational orientations affect the quality of performance. Pupils who received comments improved their Session 3 performance primarily on those components of each task which their Session 1 comments had suggested they improve. This trend is consistent with the hypothesis that task-involved, intrinsically motivated subjects will be receptive to and will actively utilise cues which can facilitate the formation of internal standards for guiding and evaluating performance (Bandura and Cervone, 1983). However, it could also be argued that they were simply "following instructions". Alternatively, since suggestions for improvement usually related to that aspect of the task for which initial scores were relatively low, their differential improvement may have been due to a ceiling effect alone. If so, enhanced performance after comments may have had little to do with task-involved or with any other kind of motivational orientation. However, were this the case one would expect to find similar increments in the performance of pupils who received both a grade and a comment. Instead, these pupils as a group did worse on Session 3 than on Session 1, and even those who improved seemed to do so without relation to the specific content of the comment they had received. Thus while our results suggest that Butler and Nisan's (1986) interpretation of similar performance increments after comments as compared with grades and no feedback in terms of higher concurrent interest alone was too narrow, they also suggest that motivational factors importantly determine the conditions under which specific guidelines will indeed be utilised to improve performance. However, further research is necessary both to establish the reliability of these rather complex effects and to pinpoint more clearly the relative effects of interest and information on subsequent performance.

While ego-involving evaluation affected interest rather differently than the extrinsic incentives employed in other studies, their effects on immediate performance seem quite similar. Thus anticipation of grades undermined performance on the novel, divergent Task B more than on the convergent Task B. In addition, while final Task A scores on Session 2 were similar after grades and comments, this was because the former generated more short words, but fewer long words than the latter. Similarly, fluency was less undermined by grades than were the other components of divergent thinking. Thus anticipation of grades, as of rewards, seems to encourage even high achievers to seek reward attainment in the easiest way possible, by concentrating on the quantity rather than on the quality of performance (Greene and Lepper, 1974).

The reduced interest and performance of low achievers who anticipated further grades, given either with or without a comment, is hardly surprising in view of the considerable anecdotal and research evidence that competitive, ego-involving situations have more adverse effects on the learning and motivation of low than of high ability students (Holt, 1964; Covington and Beery, 1976; Ames, 1984; Nicholls, 1984). Indeed, critics of grading tend to focus on the plight of the failing student. However, the present findings suggest that ego-involving settings may exact a not inconsiderable price among successful students too. Narrow preoccupation with grade attainment seems to affect the quality, if not the quantity, of immediate task performance, and to undermine divergent thinking in particular. Perhaps more significantly, the maintenance of ego-involved motivation seems to depend on the continued availability of opportunities for social comparison, so that both effort and interest wane when these are withdrawn. Thus it is suggestive that while low achievers in the grades groups expressed less interest when they anticipated further grades than when they did not, high achievers expressed least interest at post-test. In contrast, the findings for comments suggest that task-involvement is beneficial not only at low, but also at high levels of achievement.

In this context, one should note that while pupils at both levels of ability received similar comments, grades were rewarded objectively so that high achievers tended to receive high grades and low achievers low ones. Thus any effects of generalised perceptions of ability were confounded with those to be expected after experiencing specific success or failure outcomes. However, it is probably more difficult to separate these effects than is generally realised. Frieze *et al.* (1983) review several findings which suggest that objectively similar information about outcomes will be interpreted differently by different subjects, possibly as a function of such achievement-related variables as causal attributions for outcomes, achievement motivation or perceived ability. Similarly, balancing designs by giving both high and low achievers either success or failure feedback also provides some subjects with information that disconfirms expectancy, as when low achievers receive success feedback or high achievers failure feedback, which seems in itself to have considerable and differential motivational impact (Harackiewicz *et al.,* 1985). The present design, while problematic, seemed to reduce at least this pitfall, since the grades received were probably consistent with subjects' experience and expectancies, as in classroom situations.

To conclude, on a theoretical level, the results of the present study suggest that some of the difficulties faced by mean-ends analyses in conceptualising and predicting how interest and performance can be maintained or enhanced, and not just undermined, may be resolved by distinguishing not only between constrained and non-constrained but also between task-involved and ego-involved task-engagement. In addition, they suggest that different motivational orientations have implications not only for subsequent interest, but also for immediate interest and performance and for subsequent performance. Moreover, they suggest that, at least for the age-group studied, combining task and ego-involving evaluation will induce an ego-involving orientation, just as does the provision of ego-involving evaluation alone. These results seem to have implications for applied settings. While applications of the literature on rewards generally discuss when these will be more or less undermining of interest and performance (deCharms, 1983; Lepper, 1983), this study implies that promoting task-involvement may also enhance the interest and performance of most students. The comments employed here were based on a simple principle, easily applicable in the classroom, of relating both to an aspect of the task performed adequately and to one which could be improved. However, further research is clearly necessary to clarify the effects over time of systematically providing such feedback in applied settings and of reducing the use of normative grades.

Correspondence and requests for reprints should be addressed to Dr. Ruth Butler, School of Education, The Hebrew University of Jerusalem, Jerusalem 91905, Israel.

## REFERENCES

Amabile, T. M. (1979). Effects of external evaluation on artistic creativity. *J. Person. soc. Psychol.*, 37, 221-233.

Ames, C. (1984). Competitive, cooperative and individualistic goal structures: a cognitive-motivational analysis. In Ames, R., and Ames, C., (Eds.), *Research on Motivation in Education, Vol. 1,* Orlando, Flor.: Academic Press.

Bandura, A. (1982). Self efficacy mechanism in human agency. *Am Psychol.*, 37, 122-147.

Bandura, A., and Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal-systems. *J. Person. soc. Psychol.*, 45, 1017-1028.

Bandura, A., and Schunk, D. H. (1981). Cultivating competence, self-efficacy and intrinsic interest through proximal self-motivation. *J. Person. soc. Psychol.*, 41, 586-598.

Butler, R., and Nisan, M. (1986). Effects of no feedback, task-related comments and grades on intrinsic motivation and performance. *J. educ Psychol.*, 78, 210-217.

Condry, J. C. (1977). Enemies of exploration: self-initiated versus other-initiated learning. *J Person. soc. Psychol.*, 35, 459-477.

Covington, M. V., and Beery, R. (1976). *Self Worth and School Learning.* New York: Holt, Rinehart and Winston.

Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety.* San Francisco: Jossey-Bass.

deCharms, R. (1968). *Personal Causation.* New York: Academic Press.

deCharms, R. (1983). Intrinsic motivation, peer tutoring and cooperative learning: practical maxims. In Levine, J. M., and Wang, M. C. (Eds.), *Teacher and Student Perceptions.* Hillsdale, NJ: Lawrence Erlbaum.

Deci, E. L. (1975). *Intrinsic Motivation.* New York: Plenum Press.

Deci, E. L., and Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In Berkovitz, L. (Ed.), *Advances in Experimental Social Psychology, Vol. 13.* New York: Academic Press.

Frieze, I. H., Francis, W. D., and Hancuna, B. H. (1983). Defining success in classroom settings. In Levine, J. M., and Wang, M. C. (Eds.), *Teacher and Student Perceptions.* Hillsdale, NJ: Lawrence Erlbaum.

Greene, D., and Lepper, M. R. (1974). Effects of extrinsic rewards on children's subsequent intrinsic interest. *Child Dev.*, 45, 1141-1145.

Harackiewicz, J. M., Manderlink, G., and Sansone, C. (1984). Rewarding pinball wizardry: effects of evaluation and cue value on intrinsic interest. *J. Person. soc. Psychol.*, 47, 287-300.

Harackiewicz, J. M., Sansone, C., and Manderlink, G. (1985). Competence, achievement orientation and inrinsic motivation: a process analysis. *J. Person. soc. Psychol.*, 48, 493-508.

Harter, S. (1978). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. *Child Dev.*, 49, 788-799.

Holt, J. (1964). *How Children Fail.* New York: Dell.

Kruglanski, A. W., Friedman, J., and Zeevi, G. (1971). The effects of extrinsic incentives on some qualitative aspects of task performance. *J. Personality*, 39, 606-617.

Lepper, M. R. (1983). Extrinsic reward and intrinsic motivation: implications for the classroom. In Levine, J. M., and Wang, M. C. (Eds.), *Teacher and Student Perceptions* Hillsdale, NJ: Lawrence Erlbaum.

Maehr, M. L. (1976). Continuing motivation: an analysis of a seldom considered educational outcome. *Rev. educ. Res.*, 46, 443-462.

Maehr, M. L., and Stallings, W. M. (1972). Freedom from external evaluation. *Child Dev.*, 43, 177-185.

Morgan, M. (1984). Reward-induced decrements and increments in intrinsic motivation. *Rev. educ. Res.*, 54, 5-30.

Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of own attainment and the understanding that difficult tasks require more ability. *Child Dev*, 49, 800-814.

Nicholls, J. G. (1983). Conceptions of ability and achievement motivation: a theory and its implications for education. In Paris, S. G., Olson, G. M., and Stevenson, H. W. (Eds.), *Learning and Motivation in the Classroom.* Hillsdale, NJ: Lawrence Erlbaum.

Nicholls, J. G. (1984). Conceptions of ability and achievement motivation. In Ames, R. E., and Ames, C. (Eds.), *Research on Motivation in Education.* Orlando, Flor.: Academic Press.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. *J. Person, soc. Psychol.*, 43, 450-461.

Torrance, A. E., and Templeton, O. E. (1963). *Minnesota Tests for Creative Thinking.* Minneapolis: University of Minnesota Press.

Wallach, M. A., and Kogan, N. (1965). *Modes of Thinking in Young Children.* New York: Holt, Rinehart and Winston.

(*Manuscript received 2nd December,* 1986)

*Br. J. educ. Psychol.*, 58, 15-27, 1988

# FAILURE TOLERANCE AND ACADEMIC RISK-TAKING IN TEN- TO TWELVE-YEAR-OLD STUDENTS

BY MARGARET M. CLIFFORD
*(University of Iowa, USA)*

SUMMARY. The two-fold purpose of this study was to validate the School Failure Tolerance (SFT) scale and to identify developmental patterns and sex differences in children's academic risk-taking and tolerance for failure. Fourth-, fifth-, and sixth-graders were given the SFT and a chance to select and work 12 of 80 or 90 problems, of varying difficulty, in each of three content areas: mathematics, spelling and vocabulary. The SFT reliably predicted the difficulty level chosen by students at all grade levels and in all content areas. Students at all grade levels chose problems considerably below their ability levels as defined by mean performance on standardised achievement tests. This low risk-taking tendency increased markedly with grade level. Students chose problems resulting in mean absolute success of between 77 per cent and 92 per cent — levels far in excess of the 50 per cent level theoretically associated with optimum motivation. Furthermore, mean scores on two of three SFT subscales decreased significantly from grade four to grade six. There were no sex by grade interactions and only one of ten variables revealed a significant sex difference. These findings are interpreted and discussed in terms of motivational theories and classroom factors.

## INTRODUCTION

MODERATE risk-taking is a central component of many motivational theories. Atkinson's (1957, 1964) achievement motivation theory, Deci's (1975; Deci and Porac, 1978) cognitive evaluation theory, and Csikszentmihalyi's (1975, 1978) emergent motivation theory explicitly identify moderate risk-taking as an essential characteristic of the highly motivated or intrinsically motivated individual.

Many psychologists (e.g., Weiner *et al.*, 1971; Schneider, 1973; Meyer *et al.*, 1976) contend that a major reason for an individual's preference for moderately difficult tasks is the personal, skill-related information such tasks provide. In addition, it has been shown that achievement-oriented people select an easy or intermediate task to begin with, and choose tasks of progressively greater difficulty whenever they experience success (Jopt, 1974; Kuhl and Blankenship, 1979; Schneider and Posse, 1982). These findings suggest that the achievement-oriented individual is interested in, (a) identifying his highest level of competency, (b) extending that level, and (c) minimising failure.

But to engage freely in and persevere at optimally challenging tasks, one must also have a tolerance for failure or error-making. Such tolerance is likely to influence one's level of risk-taking as well as one's response to failure. Consistent with this speculation, the theory of constructive failure (Clifford, 1984) predicts that moderate risk-taking or a preference for optimally challenging tasks will be accompanied by relatively positive or constructive responses to failure (e.g., identification of the causes of failure, advice-seeking, strategy-change). Support for this prediction has been demonstrated with college students who were asked to predict the affect and behaviour of a student experiencing academic failure under varying levels of risk (Kim and Clifford, 1988).

In short, moderate risk-taking is thought to be a characteristic of individuals with (a) optimum motivation, (b) a tendency to set increasingly higher goals following success, (c) moderately high failure tolerance, and (d) a tendency to respond constructively to failure outcomes. It follows that a tolerance for error-making or failure may be a good predictor of achievement and risk-taking. Yet, we know little about children's risk-taking tendencies or failure tolerance and we know virtually nothing about the nature and function of these variables as they apply to instructional settings.

While there is a widely held belief that boys are more willing to take risks than girls (Tuddenham, 1951; Appenfels and Hays, 1961), evidence pertinent to this belief is far from conclusive. Some studies support this cultural stereotype (Crandall and Rabson, 1960; Kass, 1964; Ginsburg and Miller, 1982). Others offer only partial support (Slovic, 1966). And still others report a negligible sex difference or none at all (Jamieson, 1969; Kopfstein, 1973; Montgomery and Landers, 1974; Kogan and Dorros, 1978; Pankove and Kogan, 1986).

Research on developmental patterns of risk-taking also leaves much to be desired. While one study reveals a significant increase with age among 3- to 11-year-old children (Ginsburg and Miller, 1982), another study, conducted with 6- to 10-year-old children, failed to demonstrate age differences in risk-taking (Kass, 1964).

Evidence of children's risk-taking tendencies on school-related tasks is much more scarce than that related to social or game-like tasks. This author's review of the literature uncovered only one study of academic achievement risk-taking: Swineford (1941) reported that ninth-grade males had a higher index of test risk-taking than females. Pupils were permitted to ask for credit of two, three, or four points for any or all test items (rather than the one point traditionally assigned to each item). While they were assured that they would receive the extra credit points if the item was answered correctly, they were also told that they would lose twice the value of the item if it was answered incorrectly. The number of errors on four-point questions divided by the total number of errors plus one-half the omissions was used as a measure of risk-taking. Swineford observed that the mean risk-taking score for males was about 12 per cent higher for verbal tests and 21 per cent higher for non-verbal tests than the corresponding mean score for females.

It is ironic that academic risk-taking has received so little attention, given the widespread concern for classroom motivation and the role of risk-taking in current theories of motivation. Among the practical questions that warrant the attention of researchers are the following: when students are allowed to choose academic achievement tasks, what levels of difficulty do they prefer? Do these preferences differ depending upon whether the task involves learning or practice as opposed to testing or evaluation? Does academic risk-taking differ reliably with age and sex? How is academic risk-taking related to students' tolerance for academic error-making and failure? Can we reliably and validly measure academic risk-taking and failure tolerance? Can academic risk-taking and failure tolerance be modified, and if so, how and to what end?

The present study was designed to address a few of these questions. More specifically, the present study was used, (1) to examine the reliability and construct validity of the School Failure Tolerance (SFT) scale — a newly developed instrument intended to serve as a predictor for risk-taking in academic settings, (2) to identify sex differences and developmental patterns in failure tolerance and academic risk-taking, and (3) to examine the nature of, and relationships between children's tolerance for academic failure and their academic risk-taking behaviours — operationally defined by the difficulty level of the problems they freely choose to work.

While no predictions for sex or developmental level were made, the following construct validity hypotheses were tested for the SFT scale:

(1) School Failure Tolerance is positively correlated with academic risk-taking (i.e., difficulty level of freely-chosen problems).

(2) School Failure Tolerance is negatively correlated with accuracy of problem solutions.

(3) School Failure Tolerance subscales are positively correlated with each other.

(4) School Failure Tolerance is positively correlated with academic achievement.

## METHOD

### *Sample*

The sample for this study comprised 233 students in grades 4, 5 and 6. These students were enrolled in two separate public schools, each located in a different midwestern state. Both schools consistently average above the national median reported for the standardised achievement tests used in their respective systems. These student populations are predominantly middle-class and white. The ratio of males to the total number of subjects from grades 4, 5 and 6, respectively, were 40/83, 36/88 and 29/62.

### *School Failure Tolerance (SFT) Scale*

The SFT scale is, in some respects, similar to Kuhl's Action-Control subscale for failure-related events (Kuhl, 1985). Both measures attempt to identify, through self-report, the extent to which an individual responds constructively to failures or misfortunes. However, while Kuhl's scale focuses on failure in general, yields a cognition and behaviour score, and was designed primarily for adults, the SFT scale focuses on school failure, yields an action (i.e., behaviour), affect, and preferred task-difficulty score, and was designed primarily for 7- to 17-year-olds.

### *Academic Risk-Taking (ART) Measure*

The ART measure consisted of a booklet including two pages each of mathematics, spelling, and vocabulary items drawn from recently retired Iowa Tests of Basic Skills (ITBS) (Hieronymus and Lindquist, 1971a, 1971b; Hieronymus *et al.,* 1978a, 1978b, 1979a, 1979b) and Iowa Tests of Educational Development (ITED) (Feldt *et al.,* 1979, 1983). Each of these six problem pages contained either eight or nine rows of problems with five problems per row, for a total of 40 or 45 problems per page. All problems were multiple-choice in format.

The problems were selected and arranged such that each row increased in difficulty as measured by national norms derived from an autumn administration of these ITBS and ITED items. In general, the item difficulty for problems within a given row, as measured by percent of the norming sample that passed the item, was moderate and varied no more than 10 percentage points (e.g., 45 per cent to 55 per cent). (This meant that, if administered in the autumn of the year, the items in each row could be correctly answered by about half of the students at the grade level they represented.) The item difficulty for problems in adjacent rows differed by approximately one year, with the first row consisting of second-grade problems and the last row generally consisting of ninth-grade problems. However, since the ITED does not contain ninth-grade mathematics computation problems, the last two rows of problems on the ART mathematics pages were drawn from the eighth-grade ITED, with the second of these two rows representing a higher level of item difficult than the first. Also, since there was room for nine rows of spelling items per page, the first two rows of problems on each spelling page were drawn from the second-grade ITBS, with the second of these two rows representing a higher level of item difficulty than the first.

A major objective in designing the ART was to provide students with the maximum amount of choice in item difficulty and to ensure that a single ART measure could accommodate the desired choices of students in all three grades. Thus, as many rows of problems as possible were placed on each page. Table 1 contains the grade level for each row of problems comprising the ART measure.

The cover sheet of the ART measure contained sample problems for each of the three content areas and instructions which read as follows:

> This booklet contains MATH, SPELLING and VOCABULARY problems. Each page has many rows of problems. The first row on each page contains easy problems that are often found in 2nd grade books. The next row contains problems found in 3rd grade books. The next row contains problems found in 4th grade books, and so forth. The last row on each page contains very difficult problems. They can often be found in 8th or 9th grade books. On each page, you may choose and work only six problems. Choose *any six problems* you think you will ENJOY working. Draw a circle around the six problems that you choose to work. Some people enjoy working very easy problems. Some people like to work very difficult problems. Some people like to work problems that are not too easy or too difficult. Below is an example of some easy math, spelling and vocabulary problems. Let's work these examples together. The last page of this booklet can be torn off and used as scratch paper.

At the top of each of the six problem pages was the statement, "Choose and work the 6 MATH (SPELLING, VOCABULARY) problems you will most enjoy. Draw a circle around the 6 problems."

In addition to the scratch paper provided and attached to the back of this booklet, a page of anagrams representing 20 of the United States was appended. This page was used to occupy students who completed the ART measure relatively early.

The ART measure served as the basis for deriving two scores for each of the three content areas (i.e., mathematics, vocabulary, spelling): an ART-Difficulty score, consisting of the average row from which problems were selected and an ART-Accuracy score, indicating the number of problems out of 12 (six from each of two pages) answered correctly. The Difficulty score had a possible range of 1·16 to 8·83 for the spelling problems and 1·16 to 7·83 for the mathematics and vocabulary problems. (The difference in range resulted from the fact that the spelling task contained nine rows while the vocabulary and mathematics tasks contained eight rows per page.) The ART-Accuracy score had a possible range of 0 to 12 for each of the three content areas. It was assumed that "optimum risk-taking" could be evidenced in two ways: (1) ART-Difficulty scores representative of problems at, or just slightly above, students' achievement (i.e., grade-equivalence) level and (2) ART-Accuracy scores of approximately 6, representing 50 per cent correct responding in each content area.

### *Data gathering procedures*

The ART instrument was administered to groups of students, one grade at a time. All instructions including the same problems were read aloud by the regular classroom teacher in one school and by the experimenter in the other school. Students were explicitly reminded of the arrangement of the problems by difficulty and of their right to choose any six problems from any number of rows. Students were also reminded to draw a circle around the six problems they chose to work, to help ensure that they worked no fewer or more than six per page. Finally, students were told that they could work on the anagram task after completing their problem-solving tasks. Students were also told that they could keep the anagram sheet and

TABLE 1

GRADE LEVEL* OF PROBLEMS COMPRISING THE ACADEMIC RISK-TAKING MEASURE

| PAGE | | | | ROW | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| MATH 1 | 2nd | 3rd | 4th | 5th | 6th | 7th | 8th$^{E}$ | 8th$^{D}$ | |
| MATH 2 | 2nd | 3rd | 4th | 5th | 6th | 7th | 8th$^{E}$ | 8th$^{D}$ | |
| SPELL 1 | 2nd$^{E}$ | 2nd$^{D}$ | 3rd | 4th | 5th | 6th | 7th | 8th | 9th |
| SPELL 2 | 2nd$^{E}$ | 2nd$^{D}$ | 3rd | 4th | 5th | 6th | 7th | 8th | 9th |
| VOCAB 1 | 2nd | 3rd | 4th | 5th | 6th | 7th | 8th | 9th | |

* $^{E}$ indicates relatively easy items and $^{D}$ indicates relatively difficult items.

work it later if they did not complete it during the period. No time limit was imposed for these problem-solving tasks. All students were given time to complete the ART instrument.

One week later, the School Failure Tolerance measure was group administered to intact classes. Instructions on how to use the agree-disagree response options, which consisted of the words "yes" and "no" each printed in three different sizes, were explained: the size-of-type signalled the strength of agreement or disagreement. Sample items were worked. Students were encouraged to express their opinions honestly. Each item was read aloud by the teacher or experimenter to help ensure that students understood the items and that the time for considering the responding to each item was held relatively constant for all students.

Standardised achievement scores (ITBS) for the subjects of mathematics computation, vocabulary, and spelling, administered at mid-year were obtained only for students in one of the two schools.

## RESULTS AND DISCUSSION

### *The SFT Scale and validation hypotheses*

Responses to the School Failure Tolerance (SFT) scale were submitted to factor analysis with varimax rotation. This analysis confirmed the three anticipated failure tolerance components: Action, Preferred Difficulty, and Affect. It also provided a basis for reducing the item pool. Item retention decisions were based on a minimum factor loading of 0·40 and minimal crossloadings. The three 12-item subscales that emerged along with factor loadings for each item are presented in Table 2. The alpha coefficient for both the original set of 56 items and the reduced set of 36 items is 0·90. The alpha coefficients for the SFT subscales of Affect, Preferred Difficulty, and Action, as reported in Table 2, are 0·85, 0·88, and 0·80, respectively. Students' mean item scores (with a range of 1 to 6) for the three 12-item SFT subscales, as well as the composite SFT measure, were used for all subsequent analyses.

The first three SFT validation hypotheses were tested with data obtained from the total sample. The fourth hypothesis was tested with data obtained from one of the two schools since standardised achievement scores were unavailable for students from the other school. The alpha level for all tests was set at 0·05.

The correlations used to test the first two validation predictions are presented in Table 3. There is modest but consistent support for the first prediction: all nine of the correlations between SFT and ART-Difficulty scores are significant, and all nine of the correlations between subscale of Preferred Difficulty and the ART-Difficulty scores are significant. Furthermore, neither of the other two subscales (Action and Affect) is as strongly or consistently correlated with ART-Difficulty as is Preferred Difficulty. In summary, students' School Failure Tolerance scores, and especially their preferred Difficulty subscores, reliably predict students' choice of task difficulty, although the absolute magnitude of these coefficients tends to be small.

The strength of the relationship between failure tolerance and risk-taking might be expected to increase under two conditions not evident in the present experiment. Those conditions are the presence of performance feedback and variation in item value or "variable payoffs" (e.g., difficult items are assigned greater worth than easy items). It should be recalled that the increase in goal level (i.e., risk-taking) observed among achievement-oriented individuals following success (e.g., Jopt, 1974; Schneider and Posse, 1982; Kuhl and Blankenship, 1979) usually accompanied tasks on which they received feedback. It has also been observed that variable pay-offs induce greater risk-taking than fixed pay-offs (Ettenson and Coughlin, 1982). It is, therefore, reasonable to expect that variable pay-offs would better discriminate high and low risk-takers than fixed pay-offs (exemplified when every item or problem is worth one point regardless of its difficulty). Thus, it is reasonable to predict that failure tolerance will be a better predictor of risk-taking when feedback and variable pay-offs are present.

With regard to the second hypothesis, there is virtually no support. That is, neither SFT nor any of its subscales correlate consistently and significantly with the ART Accuracy scores. Lack of support for this prediction may, in part, be a function of students' uncertainty regarding the accuracy or correctness of their item responses. In the present study, students could more clearly judge item difficulty (accentuated by row placement) than item accuracy, since no immediate item feedback was provided. Immediate item feedback would have enabled students to control more knowingly item accuracy, and thus would have provided a fairer test of this hypothesis. It is worth noting, however, that the majority of the correlations involving ART Accuracy are negative, and thus consistent with the second prediction. In any event, it is premature to reject this hypothesis without further investigation.

A final observation relevant to the first two hypotheses concerns the confounding of ability and risk-taking behaviour. It is unclear in the present study what percentages of the ART-Difficulty and ART-Accuracy variances are attributable to student ability and what percentages are attributable to failure tolerance. If ability, as measured by standardised achievement test scores, had been controlled (e.g., used as a covariate or regression term), the predicted relationships between SFT and ART-Accuracy as well as ART-Difficulty might have been more clearly evidenced.

The third hypothesis was strongly supported. The subscales of the SFT measure were positively correlated with one another: Preferred Difficulty correlated 0·56 with Action and 0·41 with Affect, while Action and Affect correlated 0·24 with each other. These intercorrelations, considered in conjunction with the subscale reliabilities (ranging from 0·80 to 0·88), warrant the conceptualisation of School Failure Tolerance as a composite of three reasonably discrete though related factors, namely, affect, action, and preferred difficulty.

The fourth validation hypothesis received modest support. For both fourth and fifth graders, the correlation between SFT and the ITBS vocabulary subtest was positive and significant ($r = 0{\cdot}29$, $P < 0{\cdot}05$; and $r = 0{\cdot}54$, $P < 0{\cdot}01$, respectively), and the correlations between SFT and mathematics as well as spelling subtest for

TABLE 2

RELIABILITIES AND FACTOR LOADINGS FOR SCHOOL FAILURE TOLERANCE SUBSCALES

| *Subscale Items* | *Loadings* | | |
|---|---|---|---|
| | $F_1$ | $F_2$ | $F_3$ |
| *Affect (Alpha = 0·85)* | | | |
| I feel terrible when I make at mistake in school. | **0·57** | 0·08 | −0·03 |
| If I do poorly in my school work, I try not to let anyone know. | **0·48** | 0·20 | 0·01 |
| A low mark in my school work makes me feel very sad. | **0·64** | 0·02 | 0·02 |
| When I start something new in school, the first thing I think about is that I might fail. | **0·55** | 0·22 | 0·08 |
| I worry a lot about making errors in my school work. | **0·60** | 0·25 | −0·02 |
| I feel like hiding whenever I get a bad mark in school. | **0·70** | 0·04 | 0·11 |
| If I make lots of mistakes in school, I feel very moody or angry | **0·51** | 0·21 | 0·04 |
| I don't like to study with classmates because they may think I am dumb if I don't know something | **0·52** | −0·05 | 0·23 |
| When I fail at something in school, I don't like to eat, or play, or talk, or do anything. | **0·60** | 0·05 | 0·05 |
| I get very discouraged if I make errors on a task I am trying to learn. | **0·66** | 0·15 | 0·08 |
| I really dislike school work on which I make mistakes. | **0·51** | 0·31 | 0·19 |
| If I give a wrong answer to a teacher's question, I feel terrible. | **0·67** | 0·17 | 0·20 |
| *Preferred Difficulty (Alpha = 0·88)* | | | |
| I like to do school work that is difficult for me. | 0·23 | **0·63** | 0·04 |
| I would rather work problems I can do in a hurry than those that take much time and thought. | 0·08 | **0·51** | 0·35 |
| I would do almost anything to get out of working difficult problems in school. | 0·11 | **0·41** | 0·28 |
| I like to try difficult assignments even if I get some wrong. | 0·19 | **0·56** | 0·37 |
| School work that really makes me think is fun. | 0·08 | **0·56** | 0·38 |
| School work that is difficult is more fun than work that is very easy. | 0·20 | **0·72** | 0·11 |
| I would rather study a difficult course than a very easy one. | 0·16 | **0·67** | 0·13 |
| If I could chose my math problems, I would pick hard ones rather than very easy ones. | 0·06 | **0·66** | 0·30 |
| It is fun to try to answer questions that are difficult or challenging. | 0·14 | **0·54** | 0·50 |
| The easier school work is for me, the more I like it. | 0·38 | **0·58** | 0·20 |
| I like to study with classmates that enjoy working on difficult lessons. | −0·00 | **0·48** | 0·36 |
| I would rather make mistakes on a difficult task than get a perfect score on an easy but boring task | 0·08 | **0·50** | −0·12 |
| *Action (Alpha = 0·80)* | | | |
| I like to ask questions in school because I learn by asking questions. | −0·00 | 0·16 | **0·44** |
| If I can't succeed at a new school task, I give up quickly | 0·33 | −0·02 | **0·48** |
| When I make mistakes in my school work, I just keep trying and trying. | 0·03 | 0·17 | **0·64** |
| I don't like to set goals for my school work, because I might not reach them and then I feel bad. | 0·19 | 0·11 | **0·48** |
| If a school task is difficult, I try to get by without doing it. | 0·24 | 0·11 | **0·64** |
| If I do not understand something, I ask the teacher to explain it. | −0·03 | −0·05 | **0·48** |
| I would rather guess at something and get it wrong than ask a question that may sound silly. | 0·18 | 0·06 | **0·48** |
| I almost always learn a lot from the mistakes I make in my school work. | 0·09 | 0·21 | **0·56** |
| If I get a low grade in my school work, I study my errors and rework the problems I get wrong. | −0·04 | 0·21 | **0·62** |
| I usually study and correct the errors I make on school work, even if I don't have to. | −0·14 | 0·16 | **0·56** |
| I don't like to set goals for my school work. I just do the work and forget about it. | 0·06 | 0·16 | **0·46** |
| If I get a low score, I usually make up my mind to buckle down and study hard | −0·31 | 0·25 | **0·56** |

P4783

TABLE 3

CORRELATIONS OF VARIABLES BY GRADE

| SCALE | GRADE | DIFFICULTY |  |  | ACCURACY |  |  |
|---|---|---|---|---|---|---|---|
|  |  | Math. | Spell. | Vocab | Math. | Spell. | Vocab. |
| SFT | 4 | 0·29** | 0·19* | 0·19* | −0·08 | 0·04 | 0·02 |
|  | 5 | 0·37*** | 0·41*** | 0·48*** | −0·12 | 0·05 | −0·14 |
|  | 6 | 0·41*** | 0·25* | 0·37** | −0·04 | −0·01 | −0·22* |
| ACTION | 4 | 0·07 | 0·08 | 0·02 | 0·05 | 0·00 | −0·03 |
|  | 5 | 0·18* | 0·21* | 0·27** | −0·10 | −0·09 | −0·17 |
|  | 6 | 0·19 | 0·18 | 0·15 | 0·03 | −0·10 | −0·15 |
| PREF. DIFF. | 4 | 0·35*** | 0·24* | 0·23* | −0·13 | 0·07 | 0·03 |
|  | 5 | 0·37*** | 0·29** | 0·41*** | −0·14 | 0·02 | −0·16 |
|  | 6 | 0·42*** | 0·30** | 0·38*** | −0·03 | −0·14 | −0·28* |
| AFFECT | 4 | 0·22* | 0·12 | 0·18* | −0·08 | 0·00 | 0·03 |
|  | 5 | 0·32*** | 0·45*** | 0·44*** | −0·05 | −0·07 | −0·01 |
|  | 6 | 0·26* | 0·06 | 0·24* | −0·09 | 0·23* | −0·04 |

* $P < 0\cdot05$ ** $P < 0\cdot01$ *** $P < 0\cdot001$

both grades were all positive (ranging from 0·20 to 0·28) and approached significance ($P < 0\cdot15$). However, among sixth graders ($N = 26$) there was virtually no linear relationship between SFT and any of the ITBS achievement scores. Because achievement data could be obtained from less than half the total sample, it is too early to conclude whether or not failure tolerance is a reliable predictor of academic achievement, and if so, whether its predictive power has a developmental pattern. But the available data are cautiously viewed as modestly supportive of the fourth validation hypothesis.

### *Developmental trends and sex differences*

The four scores derived from the SFT as well as the six scores derived from the ART were examined using Grade × Sex ANOVAs. The means and standard deviations for these 10 variables are presented by grade level in Table 4. Preplanned comparisons which yielded values for two-tailed t-tests with 230 degrees of freedom were used to clarify the nature of significant grade effects. An alpha of 0·05 was used for all preplanned comparisons.

No significant main effects or interaction were observed for either the composite SFT measure or the Affect submeasure. For Preferred Difficulty, there was only a grade effect, $F = 4\cdot10$; $P < 0\cdot05$. The preplanned comparisons revealed that fourth-graders scored higher than both fifth- and sixth-graders, while there was no difference between the mean scores obtained by the higher two grades. For Action there was both a significant grade effect, $F = 4\cdot08$; $P < 0\cdot05$, and a significant sex effect, $F = 4\cdot25$; $P < 0\cdot05$. Females had higher (more constructive) scores than males (4·71 vs. 4·51), and the preplanned comparisons again revealed that fourth-graders scored significantly higher than fifth- and sixth-graders, while the latter two did not differ from one another.

As expected, the grade effect for ART-Difficulty was significant for all three content areas (for mathematics, $F = 6\cdot66$, $P < 0\cdot001$; for spelling, $F = 6\cdot32$,

TABLE 4

MEANS AND STANDARD DEVIATIONS FOR ART AND SFT SCORES

| Grade | Score Type | ART Content Areas | | | | | |
|---|---|---|---|---|---|---|---|
| | | MATH. | | VOCAB. | | SPELL. | |
| | | M | SD | M | SD | M | SD |
| 4 | Diff. | 3·13 | 1·22 | 3·18 | 1·44 | 3·65 | 1·71 |
| | Acc. | 10·52 | 1·49 | 9·22 | 1·90 | 9·21 | 2·54 |
| 5 | Diff | 3·78 | 1·33 | 3·63 | 1·54 | 4·21 | 1·75 |
| | Acc. | 10·52 | 1·42 | 9·44 | 2·00 | 9·32 | 2·46 |
| 6 | Diff. | 3·77 | 1·60 | 3·96 | 1·71 | 4·68 | 1·97 |
| | Acc. | 11·03 | 1·21 | 9·69 | 1·89 | 9·57 | 1·92 |

SFT Composite and Subscores

| | SFT | | Affect | | Action | | Pref Diff. | |
|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD |
| 4 | 4·15 | 0·77 | 3·71 | 1·01 | 4·81 | 0·81 | 3·94 | 1·07 |
| 5 | 3·99 | 0·79 | 3·85 | 1·09 | 4·55 | 0·81 | 3·56 | 1·06 |
| 6 | 3·85 | 0·68 | 3·61 | 1·02 | 4·46 | 0·75 | 3·48 | 1·04 |

$P < 0{\cdot}01$; and for vocabulary, $F = 5{\cdot}21$, $P < 0{\cdot}01$). However, the preplanned comparisons revealed that while fourth-graders chose significantly easier mathematics, spelling, and vocabulary problems than sixth-graders, the mean difficulty level for the two higher grades did not differ for any of the three content areas. In addition, the mathematics and spelling problems chosen by fourth-graders were significantly easier than those chosen by fifth-graders, and the difference between the vocabulary difficulty scores for these two grades approached significance ($P < 0{\cdot}06$). The only significant ART-Accuracy results involved the mathematics measure: there was a significant Grade main effect, $F = 3{\cdot}23$, $P < 0{\cdot}05$. The preplanned comparisons indicated that both fourth- and fifth-graders were less accurate (i.e., made more errors) than sixth-graders.

In summary, there were absolutely no Grade by Sex interaction effects for any of the 10 variables examined and only one Sex main effect: namely, that observed for the Action subscale of the SFT. (For this reason, descriptive statistics for the sex factor are not reported in detail.) The significant Grade effects that were observed suggest two developmental patterns: (1) School Failure Tolerance decreases with age, and (2) while older students tend to choose more difficult tasks than younger students, the difference in task difficulty preference shrinks as the grade levels of the comparative groups increase.

### *Nature and relationship of risk-taking and failure tolerance*

What about students' absolute levels of failure tolerance and risk-taking? How do they compare with levels theorists have associated with optimum motivation? As reported in Table 4, the composite SFT means for grades 4, 5 and 6 are 4·15, 3·99 and 3·85, respectively — scores only modestly above the scale midpoint of 3·5, which represents the beginning of constructive or tolerant responses to failure (in contrast to non-constructive responses or intolerance for failure). In addition, the percentage of students at these respective grade levels that had composite SFT scores below 3·5 were found to be 17 per cent, 24 per cent and 34 per cent. Thus, these data suggest that grade school students have relatively low failure tolerance in addition to the fact that failure tolerance decreases with grade.

It is also worth noting that the SFT Action means are consistently and markedly higher (approximately one standard deviation) than the SFT Preferred Difficulty means at each grade level. This suggests that, although students indicate that they respond constructively to failure, they seem intent on protecting themselves from error-making and failure by avoiding difficult tasks, whenever possible. This low risk-taking tendency is clearly exemplified in the ART data.

A comparison of the mean ART-Difficulty scores (i.e., mean row of problems selected), contained in Table 4, and the grade level associated with each row of problems, contained in Table 1, reveals two important facts: first, students tend to select items that are considerably below their ability level; and second, this item-ability discrepancy increases with grade. More specifically, note that sixth-graders have a mean ART-Difficulty score of 3·77 for the mathematics problems. But even a score of 4·0 would only indicate a selection of problems comparable in mean difficulty to those found in the 4th row — problems of average difficulty for *fifth-graders* in the *fall* of the year. Given that this study was conducted in spring, and given that subjects were from school systems that repeatedly score above the national norms on standardised achievement tests, it can be said that the sixth-graders as a group chose mathematics problems that were at least one and a half years below their ability level. Fourth-graders chose problems that were about six to eight months below their average ability level, while fifth-graders, on the average chose problems nearly a year below their ability level. Not one ART-Difficulty mean in Table 4 reflects a choice of problems that would represent truly challenging, "grade-appropriate" work. This pattern clearly supports the conclusions that academic risk-taking not only decreases with grade, but is also very low as early as fourth grade.

To characterise further the risk-taking tendencies of these students, one might examine the MTA-Accuracy scores in Table 4. These scores range from 9·21 to 11·03, with 12 representing perfect or errorless performance. These mean Accuracy scores for each grade and each content area have also been plotted in Figure 1, along with the approximate percentage of success associated with them. It can be seen that all group means reflect the selection of problems associated with mean success levels in excess of 77 per cent — levels considerably above the 50 per cent level of success theoretically associated with optimally challenging tasks and optimum motivation.

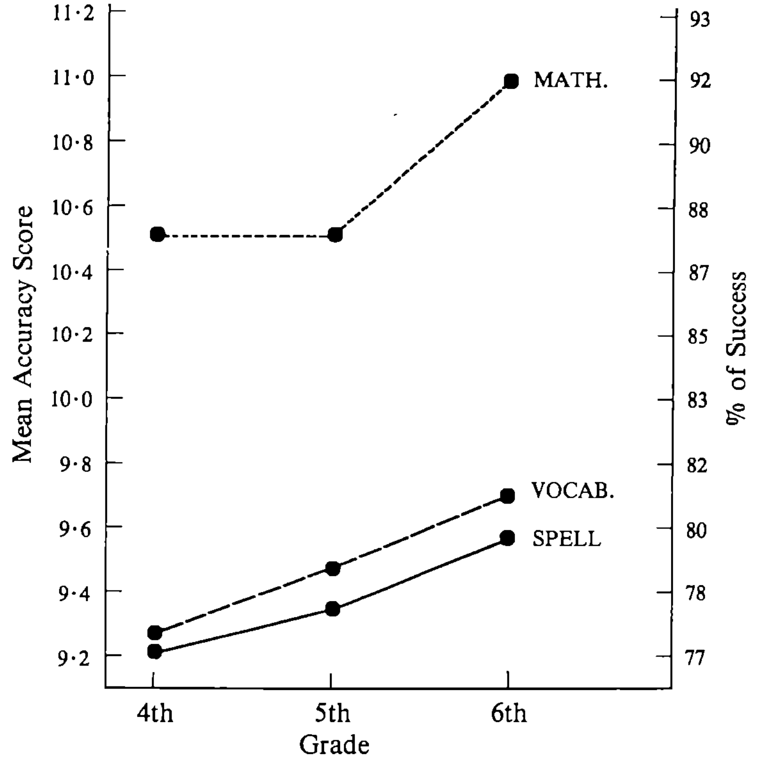
Immediate item feedback coupled with variable pay-offs might have altered these ART-Accuracy and ART-Difficulty means: knowledge of correct responses would perhaps have encouraged students to choose more difficult, higher pay-off problems as a means of acquiring increased knowledge about their ability. In turn, ART-Accuracy scores may have been lowered. But in the absence of immediate item feedback and variable pay-offs — a situation common in schools where test items and problems are frequently weighted equally and feedback is received several days after task completion — it would appear that students are intent on insuring themselves against failure. In summary, the present data suggest that grade-school students have relatively low failure tolerance and low academic risk-taking tendencies, and that both of these characteristics become more pronounced with each grade.

## CONCLUSIONS

Despite the fact that this study involved a limited sample of students from only three grades representing two schools, the following tentative conclusions would seem justified. It would appear that the SFT scale shows promise as a measure for

FIGURE 1

MEAN ART ACCURACY AND PERCENT OF SUCCESS BY GRADE LEVEL



assessing and examining failure tolerance among students. Its reliability and validity are judged satisfactory. It predicts academic risk-taking as defined by ART Difficulty scores across grade levels and content areas. It might also be expected to have increased predictive power under conditions of variable pay-offs and/or immediate item feedback. There is, however, a need to examine the usefulness of the SFT scale for students both younger and older than those participating in the present study and to examine its predictive power while controlling for ability.

Evidence of low and gradually declining failure tolerance and academic risk-taking among grade school students raises several questions. First, what factors best explain these phenomena? Even a casual observation of classrooms and instructional procedures suggests that students' avoidance of academic risk and their low tolerance for failure might be partially explained by the following: (1) reinforcements typically associated with errorless academic performance and "perfect papers", (2) the behavioural consequences of academic error-making and failure (e.g., increased assignments, extra study sessions, remedial group designations), (3) the attributional comments which often accompany error-making and imply low ability or lack of effort on the part of students (e.g., "Is that the best you can do?"; "I know you could do better if you just tried"; "I think you need some extra help"), and (4) the evaluative importance typically associated with scores obtained on most school tasks. (Scores are usually used to determine report card grades, which represent summative evaluation, rather than to provide ability information and guidance to students, which represent formative evaluation. Risk-taking and failure tolerance are likely to be lower for summative evaluation than for formative evaluation activities.)

A second question worthy of consideration is, to what extent do schools allow or invite risk-taking? Risk-taking implies choice. Academic risk-taking implies options or opportunities for students to make academic choices. To the extent that school tasks are assigned or mandated, academic risk-taking is prevented or curtailed. If, in fact, risk-taking is an essential a component of human motivation as theorists would have us believe, the presence or absence of academic risk-taking opportunities may be as crucial as any other single factor in determining and predicting the achievement orientation or intrinsic motivation of students. Academic risk-taking opportunities may take the form of test options, goal-setting options, assignment options, and even options regarding instructional materials. An increase in such options as well as means for assessing academic risk-taking opportunities would seem desirable.

A third and related question deserving attention is, what situational factors can be used to modify students' academic risk-taking tendencies? Classroom environment, teaching styles, nature of feedback, reinforcement and punishment practices are among the factors likely to affect the SFT scores and risk-taking behaviour of students. The contribution made by these and other variables should be examined, and guidelines for increasing and decreasing academic risk-taking formulated.

A final question is, what are the effects of modifying risk-taking in the classroom? It can be argued that motivation and learning theories lead to the prediction that the benefits of increasing students' academic risk-taking opportunities and tendencies will be twofold: first, motivation, as measured by such variables as self-expressed school interest, attendance, assignment completion, voluntary participation in instructional activities as well as scores on more formal measures of motivation (e.g., locus of control and need achievement scales), is likely to increase. Secondly, learning, as measured by both formative and summative evaluation — including standardised achievement scores — is likely to increase.

While lasting effects may require relatively permanent changes in the school environment and instructional procedures, short-term and task-specific effects in academic risk-taking might be generated simply by altering task-specific characteristics and instructions. For example, varying the pay-offs for ART items and encouraging students to obtain the highest possible score with their selection and completion of six problems per page would probably increase risk-taking as measured by ART-Difficulty. Modelling and reinforcing (1) the selection of difficult tasks, (2) the identification and correction of errors, and (3) moderate to high goal-setting might produce short-term and/or long-term effects in failure tolerance and academic risk-taking.

Both field and laboratory research designed to investigate the relationships between school practices, academic risk-taking, failure tolerance, motivation and learning seem warranted. This judgment is based not only on these data, but on the motivational role traditionally and theoretically assigned to risk-taking coupled with the dearth of literature pertaining to these variables.

ACKNOWLEDGMENTS —The author gratefully acknowledges the co-operation and participation of the principal, faculty and students of the Tipton Elementary School in Iowa and the Elk Mound Elementary School in Wisconsin. She also thanks William Lan, Ahyoung Kim, and Fen Chang Chou for their assistance in gathering and processing the data.

Correspondence and requests for reprints should be addressed to Professor Margaret M. Clifford, 362 L.C. College of Education, University of Iowa, Iowa City, Iowa 52242, USA.

## REFERENCES

Appenfels, E. J., and Hays, A. B. (1961). Cultural factors affecting accidents among children. *Behavioural Approaches to Accident Research.* New York: Association for the Aid of Crippled Children.

Atkinson, J. W. (1957). Motivation determinants of risk-taking behaviour. *Psychol. Rev.,* 64, 359-372.

Atkinson, J. W. (1964). *An Introduction to Motivation.* New York: Van Nostrand.

Clifford, M. M. (1984). Thoughts on a theory of constructive failure. *Educ. Psychologist,* 19, 108-120.

Crandall, V. J , and Rabson, A. (1960). Children's repetition choices in an intellectual achievement situation following success and failure. *J. genet. Psychol.,* 97, 161-168.

Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety* San Francisco: Jossey-Bass.

Csikszentmihalyi, M. (1978). Intrinsic rewards and emergent motivation. In Lepper, M. R., and Greene, D. (Eds.), *The Hidden Costs of Reward.* Hillsdale, NJ: Erlbaum Associates.

Deci, E. L. (1975) *Intrinsic Motivation.* New York· Plenum Press.

Deci, E. L., and Porac, J. (1978). Cognitive evaluation theory and the study of human motivation. In Lepper, M R., and Greene, D. (Eds.), *The Hidden Costs of Reward.* Hillsdale, NJ: Erlbaum Associates.

Ettenson, R. T., and Coughlin, R. C. (1982) Effects of type of payoff and instructions on individual risk-taking behaviour. *Psychol. Rep.,* 51, 855-860.

Feldt, L., Forsyth, R., and Lindquist, E. F. (1979). *Iowa Tests of Educational Development,* Form X-7. Chicago: Science Research Associations.

Feldt, L., Forsyth, R., and Lindquist, E. F (1983). *Iowa Tests of Educational Development,* Form Y-7, Iowa Testing Programs. University of Iowa.

Ginsburg, H. J., and Miller, S. M. (1982). Sex differences in children's risk-taking behaviour. *Child Dev ,* 53, 426-428.

Hieronymus, A. N., and Lindquist, E. F. (1971a). *Iowa Test of Basic Skills,* Form 5. Boston: Houghton Mifflin.

Hieronymus, A. N., and Lindquist, E. F. (1971b). *Iowa Test of Basic Skills,* Form 6. Boston: Houghton Mifflin,

Hieronymus, A. N., Lindquist, E. F., and Hoover, H. D. (1978a). *Iowa Test of Basic Skills,* Form 7. Boston: Houghton Mifflin.

Hieronymus, A. N., Lindquist, E. F., and Hoover, H. D. (1978b). *Iowa Test of Basic Skills,* Form 8. Boston: Houghton Mifflin.

Hieronymus, A. N., Lindquist, E. F., and Hoover, H. D. (1979a). *Iowa Test of Basic Skills,* Form 7, Boston: Houghton Mifflin.

Hieronymus, A. N., Lindquist, E. F., and Hoover, H. D. (1979b). *Iowa Test of Basic Skills,* Form 8. Boston: Houghton Mifflin.

Jamieson, B. D. (1969). The influence of birth order, family size and sex differences on risk taking behavior. *Br. J. soc. clin. Psychol.,* 8, 1-8.

Jopt, U. J. (1974). Extrinsische motivation und leistungsverhalten. Unpublished doctoral dissertation, Ruhr University, Bochum, West Germany.

Kass, N. (1964). Risk in decision-making as a function of age, sex, and probability preference. *Child Dev.,* 35, 577-582

Kim, A., and Clifford, M. M. (1988). Goal source, goal difficulty, and individual difference variables as predictors of responses to failure. *Br. J. educ. Psychol.,* 58, 28-43.

Kogan, M., and Dorros, K. (1978). Sex differences in risk taking and its attribution *Sex Roles,* 4, 755-765.

Kopfstein, D. (1973). Risk-taking behavior and cognitive style. *Child Dev.,* 44, 190-192.

Kuhl, J. (1985). Volitional mediators of cognition-behavior consistency: self-regulatory processes and action versus state orientation. In Kuhl, J., and Beckmann, J. (Eds.), *Action Control: From Cognition to Behavior.* Heidelber, New York: Springer-Verlag.

Kuhl, J., and Blankenship, V. (1979). Behavioral change in a constant environment: shift to more difficulty tasks with constant probability of success *J Person. soc. Psychol.,* 37, 551-563.

Meyer, W. U., Folkes, V. S., and Weiner, B. (1976). The perceived information value and affective consequences of choice behavior and intermediate difficulty task selection. *J. Res. Personality,* 10, 410-423.

Montgomery, G. T., and Landers, W. F (1974). Transmission of risk-taking through modeling at two age levels. *Psychol. Rep.,* 34, 1187-1196.

Pankove, E., and Kogan, N. (1968). Creative ability and risk-taking in elementary school children. *J. Personality,* 36, 420-439.

Schneider, K. (1973). *Motivation unter Erfolgrisiko.* Gottingen: Hogrefe.

Schneider, K., and Posse, N. (1982). Risk taking in achievement-oriented situations: do people really maximize affect or competence information? *Motivation and Emotion,* 6, 259-271.

Slovic, P. (1966). Risk taking in children: age and sex differences. *Child Dev.,* 37, 169-176.

Swineford, R. (1941). Analysis of a personality trait *J educ Psychol.,* 31, 438-444.

Tuddenham, R. D. (1951). Studies in reputation: I. Sex and grade differences in school children's evaluation of their peers. *Psychol Monogr.,* 66 (No 333).

Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S., and Rosenbaum, R. M. (1971). *Perceiving the Causes of Success and Failure.* New York: General Learning Press.

(*Manuscript received 6th March,* 1987)

*Br. J. educ. Psychol.*, 58, 28-43, 1988

# GOAL SOURCE, GOAL DIFFICULTY, AND INDIVIDUAL DIFFERENCE VARIABLES AS PREDICTORS OF RESPONSES TO FAILURE

By AHYOUNG KIM
*(Ewha Women's University, Seoul, Korea)*

AND MARGARET M. CLIFFORD
*(University of Iowa, USA)*

Summary. Three levels of goal difficulty (easy, moderate, difficulty) and three levels of goal source (i.e., self, other, self + other) were simultaneously examined along with two individual difference variables (i.e., mastery and failure tolerance) as predictors of responses to failure outcomes. College students majoring in education studied and gave opinions on a teacher-training activity as well as on a particular student's failure experience with this activity. This experience was described nine different ways with each student receiving only one goal source by goal difficulty combination. The strongest predictors of responses to the described failure situation were the individual difference variables; least valuable as a single predictor was goal source. Responses to failure were most constructive among students with high failure tolerance or mastery who were exposed to failure under moderate and very difficult goals. Explanations for these results as they relate to goal-setting theory and the theory of constructive failure as well as the educational implications of these results are discussed.

## INTRODUCTION

The belief that failure outcomes lead to decreased performance and negative self-perceptions has become increasingly prevalent in recent decades. As a result of this, much effort has been expended to prevent or minimise failure experiences, especially within instructional or educational settings. While learned helplessness theory and related research (Hiroto, 1974; Hiroto and Seligman, 1975) is strongly supportive of the widely held belief that failure produces devastating effects, there is also evidence derived from a variety of motivational theories, including learned helplessness, that suggests failure can also produce positive effects (Brehm, 1966; Thornton and Jacobs, 1972; Roth and Kubal, 1975; Wortman and Brehm, 1975). Furthermore, several theorists (e.g., White, 1959; Csikszentmihalyi, 1975; Deci and Porac, 1978) imply that the probability of failure *must* approximate the probability of success if intrinsic motivation and its enhancing effects are to be maximised. Thus, it becomes imperative that we more clearly delineate the conditions under which failure elicits positive and negative responses.

A recent attempt at this task is evidenced in the theory of constructive failure (Clifford, 1984). Here constructive failure is defined as, "a performance-goal ratio of less than one which produces effects such as increased persistence, task preference, task initiation, task resumption, free time devoted to task and an eventual — though not necessarily immediate — increase in task performance" (Clifford, 1984, pp. 108-109).

Based upon theories of intrinsic motivation (e.g., Csikszentmihalyi, 1975; Deci and Porac, 1978; Kruglanski, 1975; Lepper and Greene, 1978), the theory of constructive failure explicitly predicts that a goal representing optimum challenge (i.e., a moderately difficult goal) will produce more constructive responses to failure than will a non-optimum goal (i.e., a very easy or very difficult goal). In other words, the theory of constructive failure predicts a curvilinear relationship between goal difficulty and constructive responses to failure. A second prediction derived from the theory of constructive failure is that self-initiated goals produce more constructive responses to failure than imposed goals.

It could be argued that both of these goal-related predictions are inconsistent with goal-setting theory which currently predicts (1) that performance is a positive linear function of goal difficulty rather than a curvilinear function (Locke, 1967, 1968; Latham and Yukl, 1975; Locke *et al.*, 1981; Garland, 1982, 1983), and (2) with goal difficulty held constant, performance is unaffected by goal source (Latham and Saari, 1979; Latham *et al.*, 1982). Furthermore, evidence of such a linear relationship between goal difficulty and level of performance is quite plentiful.

While emphasising these goal-related distinctions between the theory of constructive failure and goal-setting theory, we hasten to add that they should not be viewed as serious conflicts. For goal-setting theory focuses on the role of goals as they affect achievement behaviour on to-be-presented tasks while the theory of constructive failure focuses on the role of goals as they affect responses to failed tasks.

In addition to examining the effects of goal source and goal difficulty on responses to failure, the present study was designed to investigate the value of individual difference variables for predicting responses to failure. There is mounting evidence that such variables are major predictors of achievement behaviour in failure situations (Diener and Dweck, 1978, 1980; Kuhl, 1981; Kuhl and Weiss, 1984; Brustein and Olbrich, 1985). The individual difference variables examined in the present study are *mastery* defined as a tendency to favour the use of relatively high evaluation criteria, and *failure tolerance,* defined as a tendency to respond to failure in a relatively constructive manner. While it is generally predicted that both mastery and failure tolerance are positively correlated with responses to failure outcomes, no specific predictions are made regarding interactions between these individual difference variables and the experimental variables of goal source and goal difficulty.

In summary, this study was designed to investigate the following: (1) whether responses to failure are a linear or curvilinear function of goal difficulty; (2) how and to what extent responses to failure are affected by goal source; and (3) what contribution the individual difference variables of mastery and failure tolerance make to explaining responses to failure. To investigate these concerns, a basic experimental design consisting of three levels of goal difficulty (i.e., difficult, moderate, easy) and three levels of goal source (i.e., self, other, self + other) was used in conjunction with the individual difference measures of mastery and failure tolerance.

## METHOD

This study consisted of presenting students with a situational description followed by a set of items used to elicit their opinions about the described character, Terry, a college student preparing for student-teaching. Steps were taken to make the situational description highly realistic and of great interest to students in an effort to maximise their personal involvement and identification with the situation.

While a situational description has the limitation of eliciting perceptions or expectations of another's behaviour and, therefore, cannot provide evidence for the ultimate concern, namely personal responses to failure, this experimental technique has three advantages: (1) it allows for the maximum and simultaneous control of variables which are otherwise difficult to control (e.g., goal source and goal difficulty); (2) it is an efficient means of gathering relatively large sets of data — a desirable criterion when one is investigating aptitude-treatment interactions; and (3) it provides a valid test of perceptions which have been found to be important determinants of behaviour and affect. Thus, despite the limitations of the situational description, this technique was judged to be appropriate for an initial investigation of the three concerns identified above.

*Sample*

The sample comprised 226 college students majoring in education at two large midwestern universities. Since the majority of these students were nearing the completion of their sophomore or junior year, it is reasonable to assume that they were anticipating their student teaching experience — an experience highly relevant to the situational description. Responses from a total of 221 students (61 males and 160 females) were used for the final analyses. Responses from another five students were discarded because they did not meet criteria set for the validation and manipulation checks which are discussed below.

*Materials*

The materials for this study consisted of an informational booklet and an experimental booklet. The informational booklet consisted of three parts. The first part provided a detailed description of the Micro-Teaching Activity (MTA), an innovative screening device supposedly being used by a major midwestern university to determine readiness for student teaching. The MTA consisted of a 45-minute instructional activity conducted by a teacher-trainee and evaluated against 17 criteria by a faculty committee. The description of the MTA included blank copies of the Request and Summary Evaluation Forms used to initiate and grade the MTA, as well as a copy of the Guidelines identifying the purpose, nature, conditions, and criteria for the MTA. (See Figure 1 for copies of the MTA Request and Summary Evaluation forms.)

Key information contained in the MTA Guidelines included the following: (1) Passing the MTA was a prerequisite for student teaching. (2) The passing score for the MTA was 68 from a possible score of 119. (3) The options available for the MTA and stipulated on the Request Form included grade level to be taught, subject content to be taught, and, most importantly, goal difficulty, familiarity/abstractness of content, and availability of prepared content materials. (See Figure 1 for the specification of these options.) (4) The MTA Request Form and the selection of options included on this form could be completed by any of three parties: the teacher-trainee's adviser, the teacher-trainee, or the teacher-trainee and adviser working jointly (representing the other, self, and self + other levels of the goal source factor, respectively).

The second part of the informational booklet consisted of a 17-item multiple-choice and true-false measure designed to ensure that students thoroughly studied and understood the MTA materials and procedures. An example of a true-false items contained in this Understanding Measure is, "The MTA Request Form *must always* be filled out by the teacher-trainee who will perform the activity". Students were encouraged to re-examine the Request and Summary Evaluation Forms and Guidelines while answering the items on the Understanding Measure.

The third part of the informational booklet consisted of a self-report measure entitled, Opinion of the Micro-Teaching Activity. This opinion sheet was actually used to obtain the individual variables of mastery and failure tolerance — both of which were situational specific measures and pertained exclusively to the MTA. The mastery measure consisted of eight statements to which students could express agreement or disagreement on a 6-point Likert scale. Examples of the mastery statements are, "I think the passing score for the MTA should be lower than 68", and "I think teacher-trainees should be evaluated against high standards, even if this means that many teacher-trainees may have to do two or more MTAs to pass". The failure tolerance measure consisted of 12 statements with a similar 6-point response format. Examples of the failure tolerance statements are, "If I were a teacher-trainee and failed the MTA by 10 or more points, I would not get very discouraged", and, "If I were a teacher-trainee and failed the MTA by 10 or more points, I would 'dig in' immediately and try to find ways to practise my teaching skills and to observe good teachers". The informational booklet was identical across all nine experimental conditions.

The experimental booklet had nine forms — one for each of the goal source by goal difficulty conditions. It contained a mock copy of a completed Request Form and Summary Evaluation Form which indicated a failure outcome (i.e., a score of 62 — 6 points below the criterion) under one of the nine experimental conditions. Figure 1 provides a copy of the forms used in the "other-goal, easy" condition. Note that the goal source manipulation (i.e., self, other, self + other) was evidenced on the Request Form under the section labelled, Source of Request, and at the end of the first line of the Summary Evaluation Form. The goal difficulty manipulation (i.e., difficult, moderate, easy) was evidenced on the Request Form as item 3 of the section labelled, Options for the Micro-Teaching Activity (MTA), and on the Summary Evaluation Form just above the names and signatures of the evaluation team.

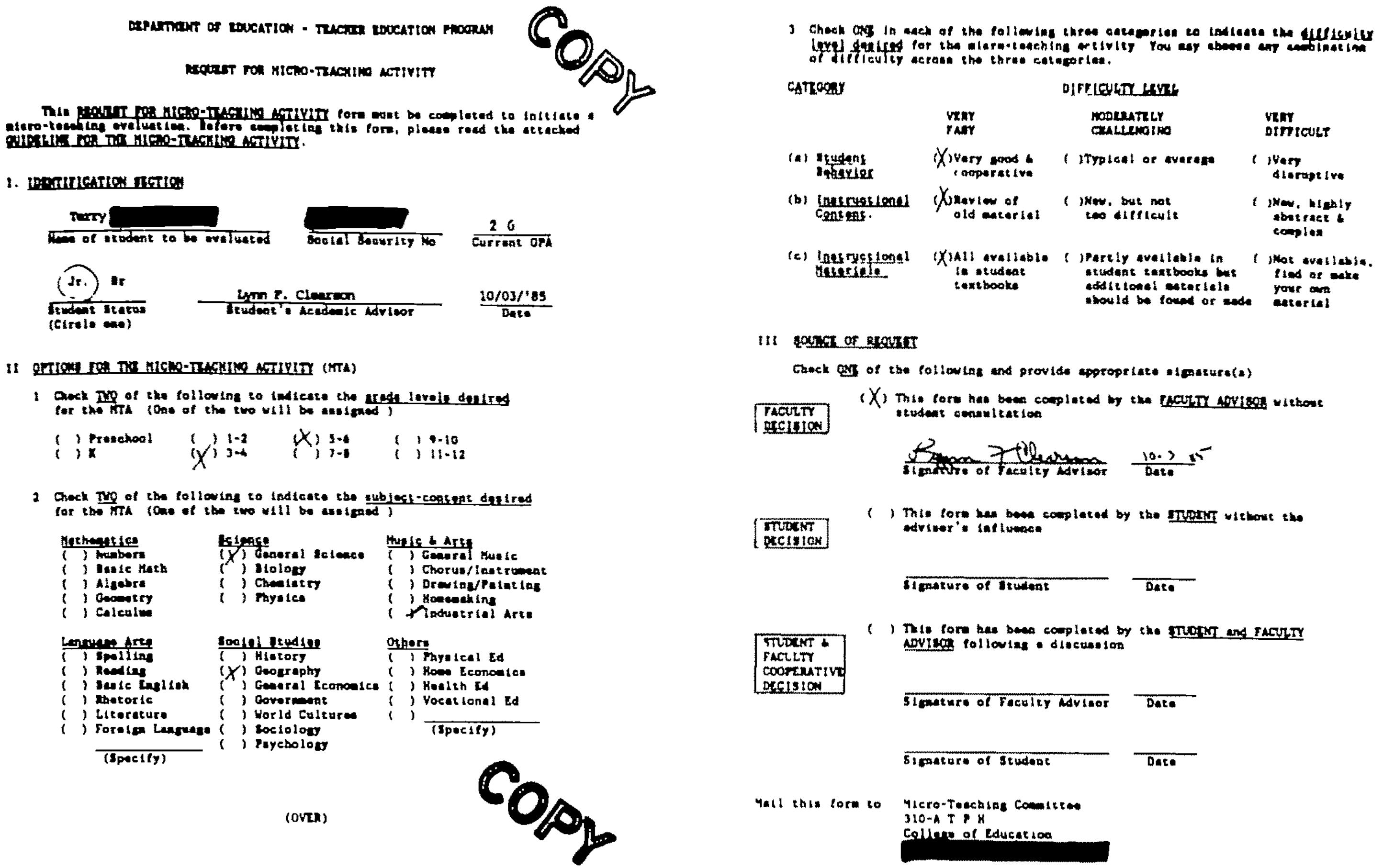
The final page of the experimental booklet contained variable check items and the two dependent variables; namely, the expected feelings and behaviour of Terry, the teacher-trainee who failed the MTA. (These dependent variables are subsequently referred to as feeling and behaviour, respectively.) The variable check items were designed to identify students who did not accurately process the source of the MTA request, the goal difficulty, or the failure outcome of the MTA. Each of the two dependent measures consisted of eight statements to which students expressed agreement or disagreement using 6-point continuums. An example of an item used to obtain the feeling measure is, "I think Terry might feel angry about this evaluation". An example of an item used to obtain the behaviour measure is, "Terry will probably seek much advice on how this MTA could have been improved". Half of the items for each dependent measure were phrased in the negative and half in the positive to control for response bias.

Two additional items were embedded in the materials to check two assumptions: a multiple-choice item was used to check the assumption that a moderate level of goal difficulty (in contrast to the easy and difficult levels) was the most appropriate level for Terry. It read, "Given that Terry has average ability as a student and teacher-trainee, I think Terry should have an MTA that is ______". The options for this item were, very easy, moderately challenging, and very difficult. A 6-point agreement-disagreement item was used to check the assumption that the "other" goal condition (i.e., having the adviser complete the Request Form without consulting the student) was plausible. This item read, "The MTA Request form can be best completed by the faculty adviser if he or she knows the teacher-trainee's abilities".

The individual difference measures (i.e., mastery and failure tolerance) as well as the dependent measures, all of which included 6-point response continuums, were scored so that high values signified favourable or desirable responses (i.e., high mastery, high failure tolerance, relatively positive feeling, constructive behaviour). Since the measures differed in the number of items comprising them, the students' average item response for each measure, which had a range of 1 to 6 and a neutral

FIGURE 1a

FRONT (LEFT) AND BACK (RIGHT) OF MTA REQUEST FORM FOR "OTHER-EASY" GOAL CONDITION



DEPARTMENT OF EDUCATION - TEACHER EDUCATION PROGRAM

COPY

REQUEST FOR MICRO-TEACHING ACTIVITY

This REQUEST FOR MICRO-TEACHING ACTIVITY form must be completed to initiate a micro-teaching evaluation. Before completing this form, please read the attached GUIDELINE FOR THE MICRO-TEACHING ACTIVITY.

I. IDENTIFICATION SECTION

Terry [redacted] — Name of student to be evaluated

[redacted] — Social Security No

2 G — Current GPA

[✓ Jr.] Sr — Student Status (Circle one)

Lynn F. Clearson — Student's Academic Advisor

10/03/'85 — Date

II OPTIONS FOR THE MICRO-TEACHING ACTIVITY (MTA)

1 Check TWO of the following to indicate the grade levels desired for the MTA (One of the two will be assigned.)

☐ Preschool
☐ K
☐ 1-2
☒ 3-4
☒ 5-6
☐ 7-8
☐ 9-10
☐ 11-12

2 Check TWO of the following to indicate the subject-content desired for the MTA (One of the two will be assigned.)

Mathematics
☐ Numbers
☐ Basic Math
☐ Algebra
☐ Geometry
☐ Calculus

Science
☒ General Science
☐ Biology
☐ Chemistry
☐ Physics

Music & Arts
☐ General Music
☐ Chorus/Instrument
☐ Drawing/Painting
☐ Homemaking
☑ Industrial Arts

Language Arts
☐ Spelling
☐ Reading
☐ Basic English
☐ Rhetoric
☐ Literature
☐ Foreign Language
(Specify)

Social Studies
☐ History
☒ Geography
☐ General Economics
☐ Government
☐ World Cultures
☐ Sociology
☐ Psychology

Others
☐ Physical Ed
☐ Home Economics
☐ Health Ed
☐ Vocational Ed
☐
(Specify)

(OVER)

COPY

3 Check ONE in each of the following three categories to indicate the difficulty level desired for the micro-teaching activity. You may choose any combination of difficulty across the three categories.

| CATEGORY | DIFFICULTY LEVEL VERY EASY | MODERATELY CHALLENGING | VERY DIFFICULT |
|---|---|---|---|
| (a) Student Behavior | ☒ Very good & cooperative | ☐ Typical or average | ☐ Very disruptive |
| (b) Instructional Content | ☒ Review of old material | ☐ New, but not too difficult | ☐ New, highly abstract & complex |
| (c) Instructional Materials | ☒ All available in student textbooks | ☐ Partly available in student textbooks but additional materials should be found or made | ☐ Not available, find or make your own material |

III SOURCE OF REQUEST

Check ONE of the following and provide appropriate signature(s)

FACULTY DECISION

☒ This form has been completed by the FACULTY ADVISOR without student consultation

Signature of Faculty Advisor

10-7 85 — Date

STUDENT DECISION

☐ This form has been completed by the STUDENT without the adviser's influence

Signature of Student — Date

STUDENT & FACULTY COOPERATIVE DECISION

☐ This form has been completed by the STUDENT and FACULTY ADVISOR following a discussion

Signature of Faculty Advisor — Date

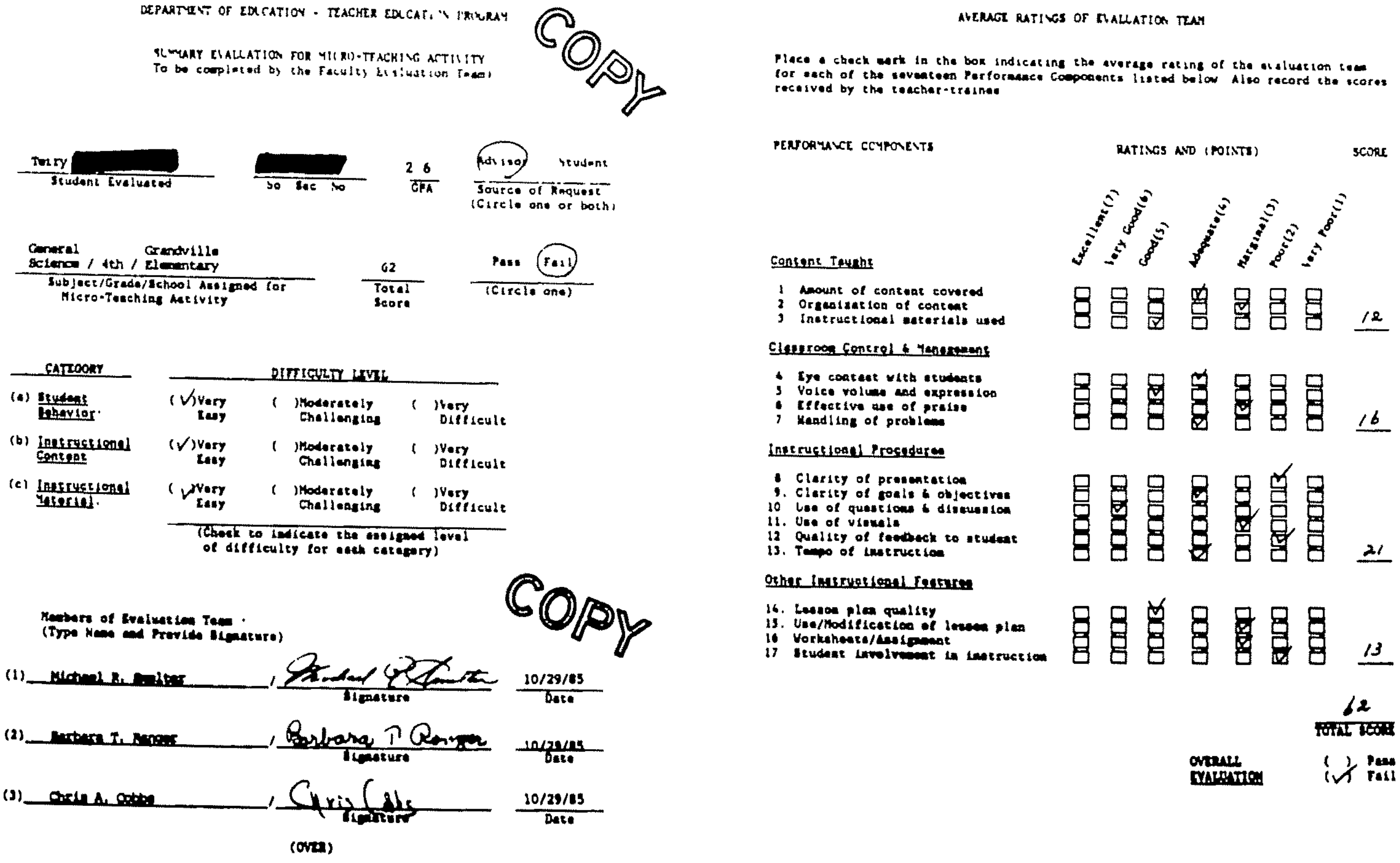
Signature of Student — Date

Mail this form to

Micro-Teaching Committee
310-A T P H
College of Education
[redacted]

FIGURE 1b

FRONT (LEFT) AND BACK (RIGHT) OF SUMMARY EVALUATION FORM FOR "OTHER-EASY" GOAL CONDITION



DEPARTMENT OF EDUCATION - TEACHER EDUCATION PROGRAM

COPY

SUMMARY EVALUATION FOR MICRO-TEACHING ACTIVITY
To be completed by the Faculty Evaluation Team)

Terry [redacted]
Student Evaluated

[redacted]
So Sec No

2 6
GPA

[✓ Advisor] Student
Source of Request
(Circle one or both)

General Science / 4th / Grandville Elementary
Subject/Grade/School Assigned for Micro-Teaching Activity

62
Total Score

Pass [✓ Fail]
(Circle one)

| CATEGORY | DIFFICULTY LEVEL | | |
|---|---|---|---|
| (a) Student Behavior | (✓)Very Easy | ( )Moderately Challenging | ( )Very Difficult |
| (b) Instructional Content | (✓)Very Easy | ( )Moderately Challenging | ( )Very Difficult |
| (c) Instructional Material | (✓)Very Easy | ( )Moderately Challenging | ( )Very Difficult |

(Check to indicate the assigned level of difficulty for each category)

COPY

Members of Evaluation Team ·
(Type Name and Provide Signature)

(1) Michael R. Smalter / Michael R Smalter 10/29/85
Signature Date

(2) Barbara T. Ranger / Barbara T Ranger 10/29/85
Signature Date

(3) Chris A. Cobbs / Chris Cobbs 10/29/85
Signature Date

(OVER)

AVERAGE RATINGS OF EVALUATION TEAM

Place a check mark in the box indicating the average rating of the evaluation team for each of the seventeen Performance Components listed below. Also record the scores received by the teacher-trainee

| PERFORMANCE COMPONENTS | Excellent(7) | Very Good(6) | Good(5) | Adequate(4) | Marginal(3) | Poor(2) | Very Poor(1) | SCORE |
|---|---|---|---|---|---|---|---|---|
| **Content Taught** | | | | | | | | |
| 1 Amount of content covered | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | |
| 2 Organization of content | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | |
| 3 Instructional materials used | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | ☐ | 12 |
| **Classroom Control & Management** | | | | | | | | |
| 4 Eye contact with students | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | |
| 5 Voice volume and expression | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | ☐ | |
| 6 Effective use of praise | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | |
| 7 Handling of problems | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | 16 |
| **Instructional Procedures** | | | | | | | | |
| 8 Clarity of presentation | ☐ | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | |
| 9. Clarity of goals & objectives | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | |
| 10 Use of questions & discussion | ☐ | ☑ | ☐ | ☐ | ☐ | ☐ | ☐ | |
| 11. Use of visuals | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | |
| 12 Quality of feedback to student | ☐ | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | |
| 13. Tempo of instruction | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | 21 |
| **Other Instructional Features** | | | | | | | | |
| 14. Lesson plan quality | ☐ | ☐ | ☑ | ☐ | ☐ | ☐ | ☐ | |
| 15. Use/Modification of lesson plan | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | |
| 16 Worksheets/Assignment | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | ☐ | |
| 17 Student involvement in instruction | ☐ | ☐ | ☐ | ☐ | ☐ | ☑ | ☐ | 13 |

62
TOTAL SCORE

OVERALL EVALUATION
( ) Pass
(✓) Fail

point of 3·5, was used for data analyses rather than the raw scores. This standardising of scores facilitated the contrast and comparison of measures and simplified the interpretation of scores.

### *Procedure*

Materials were packaged in envelopes which were handled to students as they entered the experimental room. Students were instructed to first read, study, and complete Booklet 1 (i.e., informational booklet) and then Booklet 2 (i.e., experimental booklet). Envelopes were distributed so that every ninth student was assigned to the same experimental condition. Students worked independently and at their own rate; they were encouraged to signal for assistance if they had questions.

## RESULTS

### *Manipulation and variable checks*

No student gave an incorrect answer to the failure outcome or the goal difficulty manipulation items. However, five students failed the goal source manipulation item; their data were not included in the analysis. The item used to check the plausibility of the other goal source condition produced a mean (3·22) and standard deviation (1·45), similar to those obtained for many of the other 6-point continuum items. This was taken as evidence that the students did not view this condition as highly unreasonable. As for the assumption that the moderately difficult goal was the most appropriate or "optimal" for Terry, 83 per cent of the students indicated that it was the goal Terry "should have". Of the students not choosing the moderately difficult goal as the most appropriate, 63 per cent had been assigned to the easy condition and specified, after processing Terry's failure, that an easy goal level would be most appropriate. Thus, the manipulation and control of variables critical to this study were judged to be satisfactory.

### *Descriptive statistics*

The sample means and standard deviations for the individual difference measures and the dependent measures, as well as the intercorrelations among these measures and Cronbach's reliability coefficient Alpha for each instrument, are presented in Table 1. The correlation coefficients reported are of two kinds: for the intercorrelations between the individual difference measures, which were collected before the experimental variables were manipulated, correlations were computed across all students. For the intercorrelations among the dependent measures and the correlations between individual difference measures and dependent measures, pooled within-group correlations were computed to take into consideration treatment main effects. However, such pooling is inappropriate when variables produce interactions. In such cases, the individual within-group regression slopes should be examined. (See Figures 2 through 4.)

The reliability estimates for the individual difference measures, failure tolerance and mastery, are 0·79 and 0·77, respectively. These instruments correlate 0·41 with each other. The reliability estimates for the dependent measures, feeling and behaviour, are 0·68 and 0·78 respectively. These measures have a correlation of 0·58. The correlations between the individual difference measures and the dependent measures range from 0·22 to 0·46. These data provide partial validation for both the individual difference measures and the dependent measures.

The means and standard deviations for the two dependent measures and the two individual difference measures for each of the nine treatment conditions are presented in Table 2. Given that the overall means for feeling and behaviour were 3·23 and 3·99, respectively, and given that the midpoint or neutral point of these scales was 3·5, one can conclude that the overall behavioural response to failure tended to be positive or constructive and that the overall feeling response to failure was only slightly negative.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF THE INDIVIDUAL DIFFERENCE AND DEPENDENT MEASURES, INTERCORRELATIONS, AND CRONBACH'S ALPHAS

| Measures | Mean | SD | M | FT | Feeling | Behaviour |
|---|---|---|---|---|---|---|
| Individual Difference | | | | | | |
| Mastery (M) (8 items) | 3·87 | 0·69 | 0·77* | 0·41 | 0·32 | 0·22 |
| Failure Tolerance (FT) (12 items) | 4·10 | 0·62 | | 0·79* | 0·46 | 0·40 |
| Dependent Measures | | | | | | |
| Feeling (8 items) | 3·23 | 0·61 | | | 0·68* | 0·58 |
| Behaviour (8 items) | 3·99 | 0·63 | | | | 0·78* |

*Notes.* (a) * indicates the Reliability Coefficient Alpha.
(b) For all measures, scores range from 1 to 6 and represent mean item responses.

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR INDIVIDUAL DIFFERENCE AND DEPENDENT MEASURES IN NINE GOAL SOURCE BY GOAL DIFFICULTY CONDITIONS

| | Other | | | Self + Other | | | Self | | |
|---|---|---|---|---|---|---|---|---|---|
| N = | Easy (26) | Mod. (22) | Diff. (26) | Easy (25) | Mod. (25) | Diff. (25) | Easy (24) | Mod. (25) | Diff. (23) |
| Mastery | 3·84 | 3·99 | 4·08 | 3·60 | 3·89 | 3·74 | 3·99 | 3·80 | 3·89 |
| | (0·70) | (0·61) | (0·44) | (0·73) | (0·98) | (0·61) | (0·71) | (0·65) | (0·68) |
| Failure Tolerance | 4·04 | 4·24 | 4·06 | 4·19 | 4·04 | 4·12 | 4·21 | 4·08 | 3·97 |
| | (0·69) | (0·76) | (0·64) | (0·55) | (0·68) | (0·57) | (0·51) | (0·53) | (0·69) |
| Feeling | 3·26 | 3·46 | 2·85 | 3·28 | 3·23 | 3·34 | 3·14 | 3·30 | 3·24 |
| | (0·61) | (0·73) | (0·52) | (0·46) | (0·71) | (0·62) | (0·62) | (0·44) | (0·62) |
| Behaviour | 3·78 | 4·18 | 3·99 | 3·93 | 4·11 | 4·15 | 3·82 | 4·03 | 3·98 |
| | (0·67) | (0·55) | (0·50) | (0·68) | (0·65) | (0·57) | (0·90) | (0·49) | (0·56) |

### *Regression analyses*

The results of a pilot with a similar college sample (N = 92) showed no main effect or interaction effect for the sex variable. Thus, sex was not examined as a separate factor in this study. Furthermore, scatterplots and the increment of R-squares resulting when quadratic terms were included in a regression model indicated that there was little evidence of a curvilinear relationship between either of the individual difference variables and the dependent variables. Therefore, the major analyses used to examine relationships between these two sets of variables are based on the linearity assumption.

Generalised regression analyses were conducted with each of the two individual difference variables separately entered in the regression model. The order with which variables were entered for each of the regression equations was generally consistent with that recommended by Cronbach and Snow (1977): an individual difference variable, followed by the treatment variables (i.e., goal difficulty and goal source), followed by first-order interaction terms of the individual difference variable and treatment variables, followed by a second-order interaction term.

Table 3 presents the results of the two regression analyses — one for each of the individual difference variables of failure tolerance and mastery, respectively. The R-square in each of these two models reflect the increases in the explained variance as each predictor term is sequentially entered into the equation. The F-ratio, with the final residual sum of squares used as the denominator, was used to determine whether the change in explained variance was significant. As Cronbach and Snow (1977) recommended, a significance level of $P < 0{\cdot}10$ was used to test interaction between individual difference variables and treatment variables. In reporting the results, we will first focus on the individual difference variables and then on the manipulated variables.

TABLE 3

Summary of Multiple Regression Analyses on Behaviour and Feeling Scores When Mastery (M) and Failure Tolerance (FT) Separately Entered in the Model

| Variables in Regression | df | Behaviour | | Feeling | |
|---|---|---|---|---|---|
| | | $R^2$ | F | $R^2$ | F |
| Full Model | 17 | 0·215 | 3·27*** | 0·324 | 5·71*** |
| Failure Tolerance | 1 | 0·129 | 33·36*** | 0·196 | 58·82*** |
| Goal Difficulty | 2 | 0·037 | 4·72*** | 0·012 | 1·84 |
| Goal Source | 2 | 0·005 | 0·69 | 0·004 | 0·63 |
| FT × Diff. | 2 | 0·017 | 2·26 | 0·017 | 2·56* |
| FT × Source | 2 | 0·007 | 0·93 | 0·028 | 4·28** |
| Diff. × Source | 4 | 0·001 | 0·06 | 0·038 | 2·88** |
| FT × Diff. × Source | 4 | 0·018 | 1·19 | 0·027 | 2·04* |
| Residual | 203 | 0·785 | — | 0·676 | — |
| Full Model | 17 | 0·110 | 1·48* | 0·180 | 2·62*** |
| Mastery | 1 | 0·047 | 10·62*** | 0·073 | 18·05*** |
| Goal Difficulty | 2 | 0·027 | 3·08** | 0·017 | 2·09 |
| Goal Source | 2 | 0·011 | 1·25 | 0·011 | 1·38 |
| M × Difficulty | 2 | 0·016 | 1·82 | 0·012 | 1·50 |
| M × Source | 2 | 0·001 | 0·10 | 0·003 | 0·35 |
| Diff. × Source | 4 | 0·006 | 0·34 | 0·056 | 3·47*** |
| M × Diff. × Source | 4 | 0·003 | 0·18 | 0·008 | 0·49 |
| Residual | 203 | 0·890 | — | 0·820 | — |

* $P < 0{\cdot}10$ ** $P < 0{\cdot}05$ *** $P < 0{\cdot}01$

*Notes:* (a) R-square reflects contribution of the variable in the order reported.
(b) Effects for each model were tested against the residual R-square.

*Failure tolerance*

These data strongly support the prediction that failure tolerance (FT) is positively correlated with constructive responses to failure. FT, as a single factor, explained 19·6 per cent of the feeling variance, $F(1, 203) = 58{\cdot}82$, $P < 0{\cdot}0001$, and 12·9 per cent of the behaviour variance, $F(1, 203) = 33{\cdot}36$, $P < 0{\cdot}0001$. When FT was entered into the model as the last factor, it uniquely accounted for 11·5 per cent ($P < 0{\cdot}0001$) of the variance in the behaviour variable and 15·5 per cent ($P < 0{\cdot}0001$) of the variance in the feeling variable.

A significant three-way interaction of FT, goal difficulty, and goal source was also found for feeling, $F(4, 203) = 2{\cdot}04$, $P < 0{\cdot}10$; this accounted for an additional

2·7 per cent of the variance. The nature of this three-way interaction is depicted in Figure 2 and can be described as follows: in the other goal condition, FT is a powerful predictor of feeling provided the goal is easy or moderate, but it fails to predict feeling when the goal is difficult. In the self + other goal condition, FT is a significant predictor of feeling regardless of goal source. In the self goal condition, FT is a reliable predictor of feeling only when goal difficulty is moderate. To put it another way, only under moderately difficult goals is FT a reliable predictor regardless of goal source. FT is least stable as a predictor when goals are very difficult; only if such goals are set jointly does FT reliably predict feeling following failure.

A significant goal source by FT interaction for feeling also emerged, accounting for an additional 2·8 per cent of the variance in feeling, $F(2, 203) = 4{\cdot}28, P < 0{\cdot}05$. This interaction, presented in Figure 3, indicates that FT is a better predictor of feeling when goals are set by another or self + other than when they are set by self.

Finally, FT interacted with goal difficulty. The increment in variance explained by this two-way interaction, depicted in Figure 4, was 1·7 per cent, $F(2, 203) = 2{\cdot}56$, $P < 0{\cdot}10$. FT is strongest as a predictor of feeling when goal difficulty is moderate and weakest when goals are very difficult. Taken as a whole, these two and three-way interactions for feeling indicate that FT is an especially powerful predictor of feeling following failure when moderately difficult goals have been either imposed (i.e., other goals) or jointly set (i.e., self + other goals).

### *Mastery*

There was also support for the hypothesis that mastery is positively correlated with constructive responses to failure. Mastery reliably predicted feeling; it accounted for 7·3 per cent of the variance, $F(1, 203) = 18{\cdot}05, P < 0{\cdot}0001$. Mastery also reliably predicted behaviour; it accounted for 4·7 per cent of the variance, $F(1, 203) = 10{\cdot}62, P < 0{\cdot}001$. Even when mastery was entered into the model as the last factor, and its unique contribution was computed, it accounted for 2·8 per cent ($P < 0{\cdot}01$) of the variance in the behaviour measure and 6·6 per cent ($P < 0{\cdot}0001$) of the variance in the feeling measure. Mastery did not interact with either of the manipulated variables.

It should be noted that when both Failure Tolerance and Mastery were sequentially entered into the model, Mastery made no significant contribution either as a main effect or interaction factor. That is, Failure Tolerance appears to subsume Mastery, at least as these individual difference variables were defined in the present study.

### *Goal difficulty*

Goal difficulty, a manipulated variable, was a significant predictor of behaviour and accounted for a portion of the variance (between 2·7 per cent and 3·7 per cent) not accounted for by the individual difference variables. (See Table 3 for detailed statistics on the two multiple regression analyses.) This linear relationship between goal difficulty and behaviour would seem to support goal-setting theory. However, the following should be pointed out:

(1) This simple linear relationship exists only for behaviour and not for feeling.

(2) Means for the behaviour measure for the easy, moderately difficult and difficult conditions across the goal source factor are 3·84, 4·10, and 4·04, respectively. A simple effects test showed that behaviour under moderately difficult goals was significantly more constructive than under easy goals ($P < 0{\cdot}05$), and slightly, though not significantly, better than under very difficult goals. Thus, it can

FIGURE 2

Feeling as a Function of Goal Difficulty, Goal Source, and Failure Tolerance (Unstandardised Within-Group Regression Slopes Are Shown in Parentheses With Significance Level. * $P < 0{\cdot}10$ ** $P < 0{\cdot}05$ *** $P < 0{\cdot}01$)

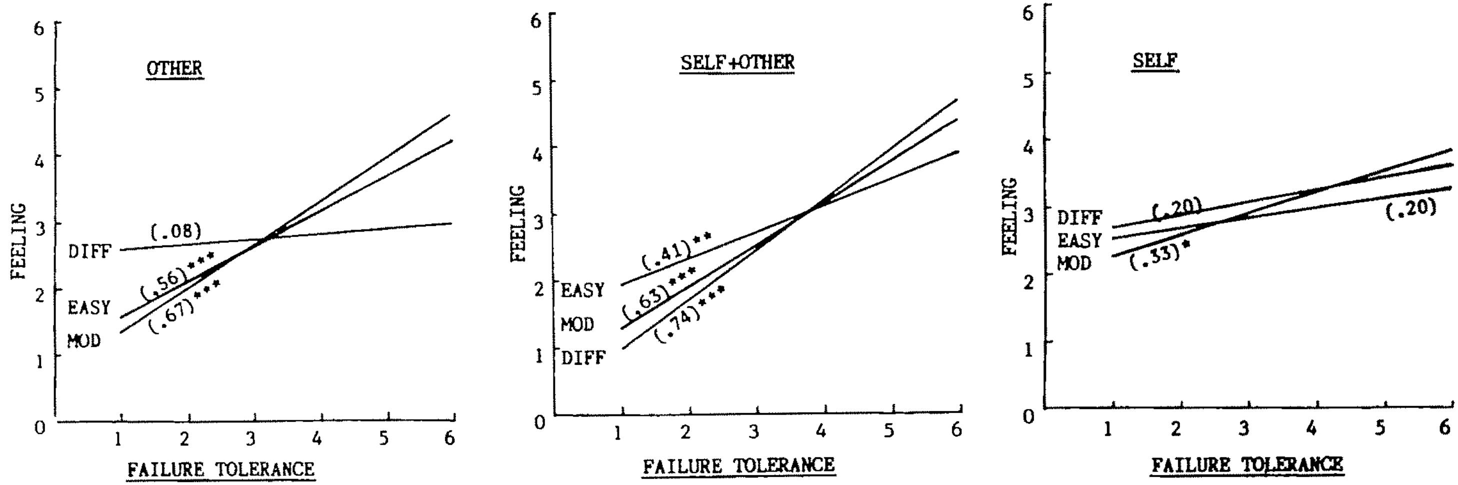


FIGURE 3

FEELING AS A FUNCTION OF GOAL SOURCE AND FAILURE TOLERANCE (UNSTANDARDISED WITHIN-GROUP REGRESSION SLOPES ARE SHOWN IN PARENTHESES WITH SIGNIFICANCE LEVEL: * P < 0·10 ** P < 0·05 *** P < 0·01)

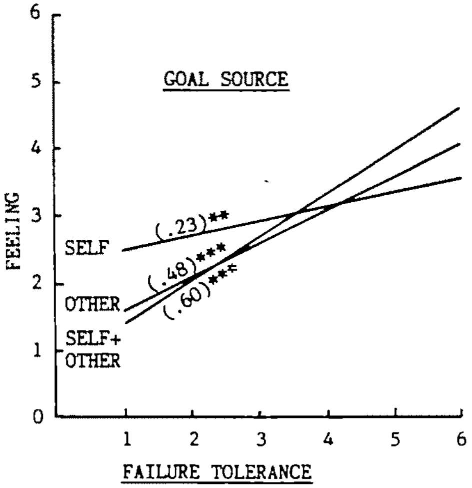
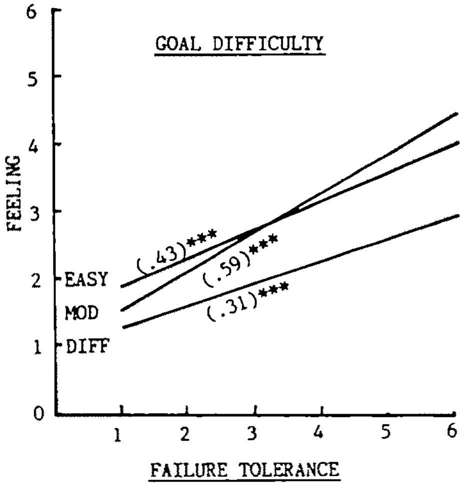


FIGURE 4

FEELING AS A FUNCTION OF GOAL DIFFICULTY AND FAILURE TOLERANCE (UNSTANDARDISED WITHIN-GROUP REGRESSION SLOPES ARE SHOWN IN PARENTHESES WITH SIGNIFICANCE LEVEL: * P < 0·10 ** P < 0·05 *** P < 0·01)



also be argued that behaviour is as much a curvilinear function of goal difficulty as it is a linear function of this manipulated variable.

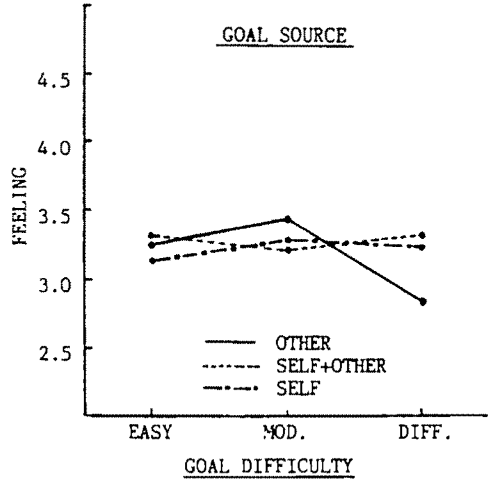
(3) In both the other and self conditions, feeling as well as behaviour produced a curvilinear trend (rather than a linear trend) as a function of goal difficulty.

(4) The predictive power of goal difficulty is noticeably weaker than the predictive power of either of the individual difference variables. Even when goal difficulty was entered into the model as the initial factor, it accounted for less than either of the individual difference variables subsequently entered into the model.

Finally there was a significant goal difficulty by goal source interaction for feeling, as depicted in Figure 5. It is best summarised by saying that goal difficulty has little or no effect on feeling, except when very difficult goals are assigned or imposed by another, in which case feeling tends to be more negative.

FIGURE 5

FEELING MEANS BY GOAL DIFFICULTY AND GOAL SOURCE



### *Goal source*

Goal source was not useful as a single predictor for either of the dependent variables. Even when goal source was entered in the model first, no significant main effect was found. However, goal source contributed to the prediction of feeling by interacting with goal difficulty as well as failure tolerance, as described earlier. Goal source contributed absolutely nothing to the prediction of behaviour, not even in an interaction. This pattern is supportive of current goal-setting theory which states that goal source has no effect on behaviour if goal difficulty is held constant (Latham and Saari, 1979; Latham *et al.,* 1982).

## DISCUSSION AND CONCLUSION

As for the three specific concerns investigated in this study, the following temporary conclusions can be made. First, the data do not offer unambiguous support for either the linear or curvilinear relationship between goal difficulty and responses to failure. It may be (a) that the goal difficulty conditions used in the present study did not adequately span the difficulty continuum and, therefore, did not provide an adequate test of these functions; (b) that the dependent measures used in the present study were simply not sufficiently sensitive to measure the differences in behaviour and affect produced by failure at a moderate *vs.* a difficult goal; and/or (c) that failure to obtain extremely difficult goals elicits responses that are, in fact, no better or worse than those elicited when one fails to attain moderately difficult goals. There is some evidence from goal-setting research to support this latter speculation (Locke, 1982; Locke *et al.*, 1984). It is clear, however, that failure to achieve relatively easy goals leads to less constructive responses than does failure to achieve challenging goals.

With regard to the second concern, these data seem to warrant the conclusion that goal source is virtually irrelevant as a predictor of responses to failure. While the relatively minor role played by goal source may, in part, be a result of the nature of the task or characteristics of the subjects, it is quite plausible that, once the task is freely chosen or imposed, the source of the goal associated with the task may be of no consequence. This postulation is consistent with goal-setting theory and not inconsistent with the other two theoretical frameworks of interest, for intrinsic motivation theories and the theory of constructive failure focus attention primarily on task source; they make no explicit claims about goal source. Finally, it is conceivable that imposed and jointly set goals for tasks judged to be important by the performer are perceived as indices of trust and confidence on the part of the authority figure (e.g., adviser, teacher, or experimenter) and, therefore, positively motivate the performer.

With regard to the third concern addressed by the present study, it can be concluded that individual difference variables do indeed help explain responses to failure outcomes. In fact, these data suggest that the level of failure tolerance and mastery an individual brings to a task are a more powerful determinant of feeling and behaviour than either goal source or goal difficulty. This implies that motivation theorists should expand and clarify the role of individual difference factors. More specifically, theorists should (a) identify those motivational individual difference variables most predictive of persistence and achievement in the face of obstacles, (b) develop and standardise instruments which can be used to assess motivational individual differences, and (c) identify intervention techniques likely to modify these individual difference factors (e.g., increase the failure-tolerance of a low achieving student).

Finally, these data appear to have practical implications for educational settings: first, moderately challenging goals are likely to produce more constructive effects following failure experiences than easy educational goals. Second, it may be less important to ensure that students are personally involved in goal-setting than to ensure that the goals they aspire to achieve are of a challenging nature. Third, when a very difficult goal is attempted, it is advisable that the goal be set by or at least endorsed by the student. Fourth, the educator's search for improved instructional outcomes might focus more on motivational differences among students. More specifically, teachers might (a) be more sensitive to their students' attitudes toward mastery and failure, (b) create environments in which tolerance for error-making and failure is nourished (e.g., reinforce choice of difficult tasks and minimise threat and punishment for failure), (c) encourage students to search for solutions to failure rather than to concentrate on, ignore, or deny failure outcomes, and (d) encourage the development of mastery-orientation (e.g., an emphasis on skill improvement and the attainment of standards of excellence in contrast to an emphasis on minimum competency goals).

Admittedly, this study leaves many important unanswered questions regarding the variables examined. Among the most important of these is the following: will such individual difference variables reliably predict the actual feelings and behaviour of the individual who fails as well as the feelings and behaviour expressed on the part of an observer (which was the focus of this study)? Recent data (Butler, 1986) investigating observer-performer differences in attributions suggest that a positive response to this question can be anticipated, especially when the observer and the described character have much in common. Other data indicate that these variables are reliable predictors of academic risk-taking defined in terms of the level of task difficulty freely chosen by students (Clifford, 1988).

Another question of major importance is: can mastery and failure tolerance be modified? The efforts of Kuhl and his colleagues to induce state orientation (Kuhl, 1981; Kuhl and Wassiljew, 1983; Kuhl and Weiss, 1984) and to induce action orientation (Kuhl, 1981, 1983; Kuhl and Weiss, 1984) as well as evidence derived from attribution training studies (Diener and Dweck, 1978, 1980) and deCharms' (1968, 1972) pawn-origin field studies, leads us to be very optimistic. Furthermore, there is evidence suggesting that mastery and failure tolerance is enhanced when failure is attributed to choice of strategy in contrast to low effort or lack of ability (Clifford, 1986a, 1986b). Thus, training in strategy attribution may be an effective method for increasing mastery and failure tolerance.

This represents a portion of the first author's dissertation which was completed at the University of Iowa in 1986. Requests for reprints can be sent to Dr. A. Kim at the Department of Psychology, Ewha Women's University, Seoul, Korea, or to Dr. M. M. Clifford, College of Education, University of Iowa, Iowa City, Iowa 52242.

## REFERENCES

BREHM, J. W. (1966). *A Theory of Psychological Reactance.* New York: Academic Press.

BRUNSTEIN, J. C., and OLBRICH, E. (1985). Personal helplessness and action control: analysis of achievement-related cognition, self-assessments, and performance *J. Person. soc. Psychol.*, 48, 1540-1551.

BUTLER, R. (1986) The role of generalised expectancies in determining causal attributions for success and failure in two social classes. *Br. J. educ Psychol.*, 56, 51-63.

CLIFFORD, M. M. (1984). Thoughts on a theory of constructive failure *Educ. Psychologist*, 19, 108-120.

CLIFFORD, M. M. (1986a). The comparative effects of strategy and effort attributions. *Br. J. educ. Psychol.*, 56, 75-83

CLIFFORD, M. M. (1986b). The effects of ability, strategy, and effort attributions for educational, business, and athletic failure. *Br. J. educ Psychol.*, 56, 169-179.

CLIFFORD, M. M. (1988). Failure tolerance and risk taking in ten- to twelve-year-old students. *Br. J educ. Psychol.*, 58, 15-27.

CRONBACH, L. J., and SNOW, R. E (1977). *Aptitudes and Instructional Methods.* New York. Irvington.

CSIKSZENTMIHALYI, M (1975). *Beyond Boredom and Anxiety.* San Francisco: Jossey-Bass.

DECHARMS, R. (1968). *Personal Causation: The Internal Affective Determinants of Behavior.* New York: Academic Press.

DECHARMS, R. (1972). Personal causation training in the schools. *J. appl soc Psychol.*, 2, 95-113.

DECI, E. L., and PORAC, J. (1978). Cognitive evaluation theory and the study of human motivation. In LEPPER, M. R., and GREENE, D. (Eds.), *The Hidden Costs of Reward.* Hillsdale, NJ: Erlbaum.

DIENER, C. I., and DWECK, C. S. (1978) An analysis of learned helplessness: continuous changes in performance, strategy, and achievement cognitions following failure. *J. Person. soc Psychol*, 36, 451-462.

DIENER, C. I., and DWECK, C. S. (1980). An analysis of learned helplessness: II The processing of success. *J. Person. soc. Psychol.*, 39, 940-952.

GARLAND, H. (1982). Goal levels and task performance: a compelling replication of some compelling results. *J. appl. Psychol.*, 67, 245-248.

GARLAND, H. (1983). Influence of ability, assigned goals, and normative information on personal goals and performance: a challenge to the goal attainability assumption. *J. appl Psychol.*, 68, 20-30

HIROTO, D. S. (1974). Locus of control and learned helplessness. *J exp Psychol.*, 102, 187-193

HIROTO, D S., and SELIGMAN, M. E. P. (1975) Generality of learned helplessness in man. *J. Person. soc. Psychol.*, 31, 311-327.

KRUGLANSKI, A. W. (1975). The endogenous-exogenous partition in attribution theory. *Psychol. Rev.*, 82, 387-406.

KUHL, J. (1981). Motivational and functional helplessness: the moderating effect of state versus action orientation. *J. Person. soc. Psychol.*, 40, 155-170.

Kuhl, J. (1983) Motivationstheoretische Aspekte der Depressionsgenese: Der Einflu von Lageorientierung auf Schmerzempfinden, Medikamentenkonsum und Handlungs-kontrolle. In Wolfersdorf, M., Straub, R., and Hole, G. (Eds.). *Der depressiv Kranke in der psychiatrischen Klinik: Theorie und Praxis der Diagnostik und Therapie.* Regensburg: Roderer.

Kuhl, J., and Wassiljew, I. (1983). Intrinsic task-involvement coping with failure, and problem-solving: motivational and emotional determinants of the complexity of action plans. (Manuscript submitted for publication.) Max Planck Institute for Psychological Research. Munich, FRG.

Kuhl, J., and Weiss, M. (1984). Performance deficits following uncontrollable failure: impaired action control or global attributions and generalised expectancy deficits. (Manuscript submitted for publication.) Max Planck Institute for Psychological Research, Munich, FRG.

Latham, G. P., and Saari, L. M. (1979). The effects of holding goal difficulty constant on assigned and participatively set goals. *Academy of Management J.,* 22, 163-168.

Latham, G. P., Steele, T. P., and Saari, L. M. (1982). The effects of participation and goal difficulty on performance. *Personnel Psychol.,* 15, 677-686.

Latham, G. P., and Yukl, G. A. (1975). A review of research on the application of goal setting on organisations. *Academy of Management J.,* 18, 824-845.

Lepper, M. R., and Greene, D. (1978). Overjustification research and beyond: toward a means-end analysis of intrinsic motivation. In Lepper, M. R., and Greene, D. (Eds.), *The Hidden Costs of Reward.* Hillsdale, NJ: Erlbaum.

Locke, E. A. (1967). Relationship of goal level to performance level. *Psychol. Rep.,* 20, 1068.

Locke, E. A (1968). Toward a theory of task motivation and incentives. *Organ. Behav. hum. Perform.,* 51, 324-329.

Locke, E. A. (1982). Relation of goal level to performance with a short work period and multiple goal levels. *J. appl Psychol.,* 67, 512-514.

Locke, E. A., Frederick, E., Buckner, E., and Bobko, P. (1984). Effects of previously assigned goals on self-set goals and performance. *J. appl. Psychol,* 69, 694-699.

Locke, E. A., Saari, L. M., Shaw, K. N., and Latham, G. P. (1981). Goal setting and task performance: 1969-1980. *Psychol. Bull.,* 90, 125-152.

Roth, S., and Kubal, L. (1975). Effects of noncontingent reinforcement on tasks of differing importance: facilitation and learned helplessness. *J. Person. soc. Psychol.,* 32, 680-691.

Thornton, J. W., and Jacobs, P. D. (1972). The facilitating effects of prior inescapable unavoidable stress on intellectual performance. *Psychon. Sci.,* 26, 185-187.

White, R. W. (1959). Motivation reconsidered: the concept of competence. *Psychol. Rev.,* 66, 297-333.

Wortman, C. B., and Brehm, J. W. (1975). Responses to uncontrollable outcomes: an integration of reactance theory and the learned helplessness model. In Berkowitz, L., (Ed.), *Advances in Experimental Social Psychology* (Vol. 8). New York: Academic Press.

(*Manuscript received 9th February,* 1987)

*Br. J. educ. Psychol.*, 58, 44-54, 1988

# THE RELATIONSHIP BETWEEN SCHOOL TYPE AND ADOLESCENT SELF ESTEEM, ATTRIBUTION STYLES, AND AFFILIATION NEEDS: IMPLICATIONS FOR EDUCATIONAL OUTCOME

By ANNE E. FOON
*(Canberra College of Advanced Education, Australia)*

SUMMARY. This paper explores the notion that the type of school attended will affect adolescents' self esteem, locus-of-control orientation, and affiliation needs. It proceeds to examine the view that these are in turn linked to success in particular subject areas, assessments of achievement in these same subjects, perceptions of the importance of school life and future career options.

A sample of 1,675 school children, 896 males and 779 females from sixteen non-government schools was employed. The majority (79 per cent) attended single-sex schools while the remainder attended co-educational schools. Major results indicated that type of school attended did have consequences in terms of adolescents' stated preferences and rated achievement in subjects. Students attending co-educational school were found to have more traditional subject preferences and assessments of achievement in those subject areas than students attending single-sex schools. Students attending co-educational schools were also found to have lower self esteem, and greater affiliation with peers than students from single-sex schools. However, these relationships were not simple. Sex differences mediated the major associations reported between type of school attended and adolescent self perceptions.

Interpretations of these and other results and suggestions for further exploration of associations are given.

## INTRODUCTION

OVER the last decade or so in a number of countries a question which has continued to be evaluated is the advantages of single-sex versus co-educational schooling. The belief that satisfactory patterns of educational attainment can be achieved through the gross manipulation of school variables is debatable. Factors much less tangible than the organisational structure of the school may operate to determine the achievement of children in schools. This area would appear worthy of more detailed research. This paper explores first, the relationship between school type and adolescent self esteem, attribution styles and affiliation patterns and second, the effect these relationships have on education outcome. Educational outcome is considered here to be multidimensional in nature and includes perceptions of the importance of school life, and future career options, as well as grades and estimations of ability in particular subject areas.

There is a literature that suggests that females and males gravitate to particular subject areas and perform differentially in them because of biological processes. For the most part the evidence for the biological basis for differences in school attitudes and achievement has been discounted (Maccoby and Jacklin, 1974). The lack of evidence for an inherent difference between the sexes in choice of subject and achievement has led researchers to focus on socio-cultural factors important to educational outcome.

An example of research focusing on broad socialisation processes for outcome is Horner's (1972) work which explores the seeming incompatibility between competence and intellectual achievement and feminity. By inference, Horner argues that academic excellence is compatible with notions of masculinity. The related idea that fear of success is incompatible with high performance is an interesting hypothesis. Horner (1970) suggests from evidence provided by earlier researchers on girls' fear of success in mixed competitive settings that there may be advantages for girls in girls-only groupings. A more radical position is that advanced by Sarah *et al.* (1986). They argue that within co-educational schools the academic and social relations contribute to the subordination of females, lowering their assessment of their academic competence and ensuring lower academic performance.

If this argument is correct, males would be expected to rate their academic achievement higher than females, and to prefer traditionally male-oriented subjects (mathematics, science) more often than females. Furthermore, greater differences in estimates of achievement and subject preferences would be expected between boys and girls in co-educational schools compared to those in single-sex schools.

A second area of research evidence relates to the suggestion that the measured self esteem of adolescent females is significantly lower than that of their male counterparts (Dynan, 1980; Edgar, 1985). Secondary schooling, it is argued, has a debilitating effect on female self esteem (Gurney, 1980). It is possible that females' need for peer group acceptance found to be prevalent in adolescence (Eder, 1985) leads to poor self esteem. If these assumptions are accurate, adolescent females will have lower estimations of themselves and, related to this, higher needs for peer group acceptance than males of the same age. Following on from this, there should be no difference between the self esteem of girls in co-educational or single-sex schools. If, however, it is direct competition which causes low esteem, girls' estimations of themselves will be higher in single-sex than in co-educational schools, while boys' esteem should remain relatively stable across school type.

As just indicated, a number of researchers, most recently Eder (1985), have investigated the importance of peer affiliations and feelings of self worth for school outcomes. In particular, it has been considered that females, in contrast to males, are very dependent on the group during adolescence for their sense of well-being. From this, it could be argued that in adolescence females will have greater needs for affiliation with peers than males.

With regard to different forms of affiliation, females would be expected to have greater needs for affiliation with peers than with their family or school while the reverse pattern might reasonably be expected for males. If identification with peers and a high degree of dependence on groups is an inherent part of the socialisation process females will, irrespective of the type of school attended, score higher on both of these measures than males. If, on the other hand, females' attitudes toward peers and dependence on the group are communicated specifically in association with males, then differences should occur as a function of school type. That is, females from single-sex schools should be less dependent on the group and have less need for peer affiliation than females attending co-educational schools.

Finally, there is some evidence that indicates that females tend to attribute success in subjects to unstable factors such as an easy test on a day when they felt good while males tend to attribute success in subjects to stable factors such as ability (Wolleat *et al.*, 1980). At every level, females have been found to perceive academic achievement as being out of their control (Fennema and Sherman, 1977). Not only is this the case for evaluations of academic subjects but also there is evidence that females generally are more external in their control orientations than males, that is, females have a greater likelihood of attributing life events to external forces such as chance, luck, or powerful others (Phares, 1976).

The implications of this for academic achievement and estimations of academic worth are potentially great. It is possible that females' dominant attribution styles are structured, at least in part, in competitive situations with males; the classroom is such a situation. If this is the case, females from co-educational schools would be expected to be more external in their control orientations than females from single-sex schools. If, on the other hand, locus-of-control is directly related to sex and affects academic results and estimations of academic ability, females are likely to have poorer grades and poorer self-perceptions of academic performance than males, regardless of type of school attended.

Given the arguments above, it is also reasonable to expect that females with predominantly external orientations will have lower self esteem and a greater need for group/peer affiliation than males, who are characterised by predominantly internal orientations. Investigating the importance of school type for locus-of-control and in turn establishing the importance of that relationship for academic achievement and student evaluations of themselves and the importance of school life would seem valuable in helping clarify the direction of these associations.

The study to be reported here investigates differences in school type and adolescent self esteem, locus-of-control, and affiliation patterns as well as individuals' perceptions of current ability, self-reported academic grades, attitudes to school life and aspirations for future careers.

## METHOD

### *Independent measures*

(1) School type: schools were scored as either girls only (1), boys only (2) or co-educational (3).

(2) Self esteem: the Piers-Harris Children's Self Concept Scale (Piers, 1984) was employed to assess self esteem. The scale comprises 80 items. The method of scoring was as follows: items measuring high self esteem were scored +1 for "yes" and −1 for "no" responses, while items measuring low self esteem were scored −1 for "yes" and +1 for "no" responses.

(3) Affiliation patterns: peer, family and school affiliation scales were adapted from scales devised by Richmond (1985). Each scale comprised five items and was scored from 1 (never or almost never) to 4 (always or almost always).

(4) Locus-of-control: the short (21-item) version of Nowicki and Strickland's (1973) locus-of-control scale for children was employed to assess attribution styles. items were scored either yes (scored 2) or no (scored 1).

### *Dependent measures*

The major dependent measures were reported mark/grade in mathematics, English and science for the previous year (ranging from 5: highest grade to 1: fail grade); rank ordering of preference for mathematics, science and English (ranging from 3: most preferred to 1: least preferred); and ratings of academic performance in each of these three subjects (ranging from 5: excellent to 1: poor).

The degree of importance that students attached to their school work and the degree to which they reported their parents, teachers and peers attached importance to their school work was measured on scales ranging from "very important" (scored 5) to "not at all important" (scored 1). Also, students were asked to indicate their occupational ambition (scored from 6: low occupational ambition to 1: high occupational ambition).

*Background measures*

Father's and mother's occupation were obtained and scored, as for the occupational ambition measure, on a 6-point scale developed from a listing of the Australian workforce (Broom and Jones, 1976) with 6 representing semi-skilled and unskilled (low-status) and 1 representing professional (high-status) occupations. These measures were included to take account of the potential importance of these major background variables for educational outcomes.

## PILOT STUDY

A pilot study was conducted with 273 year 10 students from three non-government schools (one all-boys, one all-girls and one co-educational), two situated in Canberra and one in Sydney, to establish the reliability of the scales. Coefficient alphas for each of the scales were as follows:

| | |
|---|---|
| Self esteem: | 0·85 |
| Locus-of-control: | 0·65 |
| Affiliation with school: | 0·71 |
| Affiliation with peers: | 0·69 |
| Affiliation with family: | 0·82 |

A number of items were replaced in several of the scales for the main study in an attempt to achieve higher inter-item correlations: Co-efficient alphas for the main study (N = 1,675) were somewhat improved:

| | |
|---|---|
| Self esteem: | 0·91 |
| Locus-of-control: | 0·74 |
| Affiliation with school: | 0·81 |
| Affiliation with peers: | 0·81 |
| Affiliation with family: | 0·82 |

## MAIN STUDY

*Sample*

The sample for the main study comprised 1,675 schoolchildren, 779 females and 896 males. They were all year 10 students with an average age of 15 years.

The schools selected to provide students to participate in this study were consistent with regard to type of school. The reason an attempt was made to keep school type comparable was because it appears that research efforts are potentially confounded by differences within samples with regard to this variable. A good example is a recent study conducted with year 12 students in Australia (Carpenter, 1985) which employed a sample of females from non-government single-sex schools and government co-educational schools in an attempt to investigate the effect of a single sex schooling on females' academic achievement. It is conceivable that unmeasured differences between government and non-government schools in general or between co-educational government and single-sex non-government schools in particular may have confounded the results reported. In order to overcome the possible debilitating effects of these differences, the present study employed pupils from 16 non-government schools (five all-girls, six all-boys and five co-educational). All schools were located in Melbourne, Australia.

*Procedure*

Each pupil was asked to complete an 18-page questionnaire, instructions included. The questionnaire was administered to class size groups by teachers. Instructions for each scale were read out by the administrator, while pupils read along. Forty minutes were allocated for the completion of the questionnaire and times for each task were specified in a time sheet prepared by the author. Each administrator was asked to adhere closely to the time schedule.

## RESULTS

The majority of respondents reported their previous year grades in mathematics, English and science as average or above average, 50 per cent, 59 per cent and 51 per cent respectively. In terms of rated preference for subjects, science was the most preferred followed by English and mathematics: the differences in percentages were not large. The rated achievement in all subjects was high, with the majority of respondents rating their performance as average or above average; 72 per cent in the case of mathematics, 81·5 per cent in the case of English and 75 per cent for science.

Nearly all pupils rated doing well in school as important or very important to themselves and their parents, 95 per cent and 95·5 per cent respectively. There was a wider distribution of scores for the rated importance of doing well in school for teachers and peers. In the case of teachers, the proportion of students rating high performance as being important or very important was still high (78·5 per cent), but considerably less than ratings for self and parents. The importance pupils felt peers attached to their doing well in school was much less. The majority of pupils (63·5 per cent) perceived that peers considered their doing well in school only marginally important or not at all important.

The greatest proportion of pupils had high-status occupational ambitions. Fathers' occupations were also reported as being mainly in the high-status categories: professionals/managers made up 67 per cent of the sample, clerical and skilled manual 22 per cent, and semi-skilled/unskilled occupations 5 per cent. The percentage of mothers in professional/managerial occupations was reported as 24·5 per cent of the sample. The proportion of mothers in clerical and skilled manual was slightly higher (29·5 per cent of the sample), while the remainder were either employed in semi-skilled/unskilled occupations or unemployed (the majority of the latter engaged in home duties).

The greater proportion of pupils had high self esteem. High group dependence scores were also found to be prevalent as were high affiliation scores with school, peers and family. The split between internal and external locus-of-control orientations was roughly equivalent, with internal attribution styles being slightly more prevalent.[1]

Significant differences were obtained between the sample of females and the sample of males employed in this study on several measures (see Table 1). Females rated their preference for and achievement in English more favourably than males, while males rated their achievement in mathematics and science more favourably than females. Females perceived that their parents attached greater significance to their doing well in school than males, while males reported that their friends and teachers attached greater significance to their doing well at school than did females. Moreover, males reported greater affiliations with family and peers than females. The occupational ambitions of males were higher than those of females. Also, the self esteem of males was found to be significantly higher than that of females.

In order to break down these differences between males and females, a number of more specific comparisons were conducted. Males and females from co-educational schools were compared first. As for the total sample, females from co-educational schools rated their preference for and achievement in English

---

[1] A complete list of the distribution of scores on major variables is available from the author.

TABLE 1

SIGNIFICANT DIFFERENCES BETWEEN MALES (N = 896) AND FEMALES (N = 779)

| | | $\bar{X}$ | SD | t | P |
|---|---|---|---|---|---|
| (High) Rated preference for English | Males | 2·33 | 1·13 | −7·31 | 0·01 |
| | Females | 2·73 | 1·05 | | |
| (High) Rated achievement in mathematics | Males | 3·62 | 0·95 | 3·32 | 0·01 |
| | Females | 3·46 | 0·91 | | |
| (High) Rated achievement in English | Males | 3·34 | 0·83 | −2·23 | 0·05 |
| | Females | 3·43 | 0·78 | | |
| (High) Rated achievement in science | Males | 3·53 | 0·93 | 2·35 | 0·05 |
| | Females | 3·42 | 0·90 | | |
| (High) Perceived value parents attach to doing well in school | Males | 3·53 | 0·60 | −2·75 | 0·01 |
| | Females | 3·60 | 0·58 | | |
| (High) Perceived value teachers attach to doing well in school | Males | 3·14 | 0·76 | 3·89 | 0·01 |
| | Females | 2·99 | 0·83 | | |
| (High) Perceived value friends attach to doing well in school | Males | 2·31 | 0·81 | 3·30 | 0·01 |
| | Females | 2·17 | 0·80 | | |
| (High) Occupational ambition | Males | 2·56 | 1·55 | 5·88 | 0·01 |
| | Females | 2·09 | 1·38 | | |
| (High) Self esteem | Males | 1·99 | 8·19 | 3·07 | 0·01 |
| | Females | 0·64 | 8·63 | | |
| (High) Family affiliation | Males | 9·34 | 3·51 | 2·83 | 0·01 |
| | Females | 8·86 | 3·24 | | |
| (High) Peer affiliation | Males | 11·10 | 2·38 | 4·77 | 0·01 |
| | Females | 10·53 | 2·42 | | |

significantly higher than males. Also, males rated their achievement in mathematics and science higher than females. In addition, males from co-educational schools were found to express greater preferences for mathematics and science subjects than females and to aspire to higher status occupations than females. Last of all, males from co-educational schools reported greater affiliation with their peers than females from co-educational schools. Table 2 presents these results.

In the case of comparisons between males and females from single-sex schools, some similarities were obtained between these and those obtained for the total sample. Table 3 presents these results. The self esteem of males was significantly higher than those of females from single-sex schools. The importance that females perceived their parents attached to their doing well in school was again higher than for males. As for the total sample, males perceived that friends and teachers attached greater importance to their doing well in school than females. Again as for the total sample comparison, males from single-sex schools were found to have greater affiliations with family and peers than females from single-sex schools. Finally, males rated their preference for science and achievement in mathematics higher than females, as well as indicating a greater preference for science and English subjects.

Given the overlap in several of the significant differences obtained in the above

TABLE 2

SIGNIFICANT DIFFERENCES BETWEEN MALES (N = 194) AND FEMALES (N = 163) FROM CO-EDUCATIONAL SCHOOLS

| | | $\bar{X}$ | SD | t | P |
|---|---|---|---|---|---|
| (High) Rated preference for mathematics | Males | 2·55 | 1·12 | 2·64 | 0·01 |
| | Females | 2·24 | 1·08 | | |
| (High) Rated preference for English | Males | 2·03 | 1·14 | −6·94 | 0·01 |
| | Females | 2·86 | 1·05 | | |
| (High) Rated achievement in science | Males | 2·92 | 0·96 | 3·31 | 0·01 |
| | Females | 2·58 | 1·05 | | |
| (High) Rated achievement in mathematics | Males | 3·65 | 0·96 | 3·15 | 0·01 |
| | Females | 3·34 | 0·84 | | |
| (High) Rated achievement in English | Males | 3·19 | 0·91 | −3·06 | 0·01 |
| | Females | 3·45 | 0·68 | | |
| (High) Rated achievement in science | Males | 3·52 | 0·92 | 2·63 | 0·01 |
| | Females | 3·27 | 0·82 | | |
| (High) Occupational ambition | Males | 2·57 | 1·58 | 2·74 | 0·01 |
| | Females | 2·06 | 1·47 | | |
| (High) Peer affiliation | Males | 10·81 | 2·52 | 2·49 | 0·05 |
| | Females | 10·12 | 2·61 | | |

TABLE 3

SIGNIFICANT DIFFERENCES BETWEEN MALES (N = 702) AND FEMALES (N = 616) FROM SINGLE-SEX SCHOOLS

| | | $\bar{X}$ | SD | t | P |
|---|---|---|---|---|---|
| (High) Rated preference for English | Males | 2·41 | 1·12 | −4·68 | 0·01 |
| | Females | 2·69 | 1·05 | | |
| (High) Rated preference for science | Males | 2·72 | 1·02 | 3·55 | 0·01 |
| | Females | 2·51 | 1·05 | | |
| (High) Rated achievement in mathematics | Males | 3·61 | 0·95 | 2·16 | 0·05 |
| | Females | 3·50 | 0·92 | | |
| (High) Perceived value parents attach to doing well in school | Males | 3·53 | 0·61 | −2·50 | 0·05 |
| | Females | 3·61 | 0·58 | | |
| (High) Perceived value teachers attach to doing well in school | Males | 3·12 | 0·77 | 3·53 | 0·01 |
| | Females | 2·97 | 0·82 | | |
| (High) Perceived value friends attach to doing well in school | Males | 2·31 | 0·80 | 2·69 | 0·01 |
| | Females | 2·18 | 0·80 | | |
| (High) Self esteem | Males | 2·55 | 7·93 | 3·75 | 0·01 |
| | Females | 0·73 | 8·56 | | |
| (High) Family affiliation | Males | 9·36 | 3·52 | 2·56 | 0·05 |
| | Females | 8·88 | 3·22 | | |
| (High) Peer affiliation | Males | 11·17 | 2·33 | 4·07 | 0·01 |
| | Females | 10·64 | 2·36 | | |

comparisons, it was felt that it would be informative to compare subjects from co-educational and single-sex schools as total groups and then females from single-sex versus co-educational schools and males from single-sex versus co-educational schools.

Peer affiliation needs were greater for subjects from co-educational schools, while high self esteem was significantly more prevalent in single-sex schools (see Table 4).

TABLE 4

SIGNIFICANT DIFFERENCES BETWEEN SUBJECTS FROM SINGLE-SEX SCHOOLS (N = 1,318) AND CO-EDUCATIONAL (N = 357) SCHOOLS

| | | $\bar{X}$ | SD | t | P |
|---|---|---|---|---|---|
| (High)<br>Self esteem | Single<br>Co-ed | 1·68<br>0·12 | 8·28<br>8·84 | 2·78 | 0·01 |
| (High)<br>Peer affiliation | Single<br>Co-ed | 10·42<br>10·89 | 2·58<br>2·36 | −3·05 | 0·01 |

In the case of comparisons of females from single-sex versus co-educational schools, several significant differences were obtained. Females from single-sex schools were more likely to prefer science subjects and rate their achievement high in science and mathematics compared to females from co-educational schools. By contrast, females from co-educational schools rated their preference for English significantly higher than females from single-sex schools. Table 5 presents these findings:

TABLE 5

SIGNIFICANT DIFFERENCES BETWEEN FEMALES FROM SINGLE-SEX (N = 616) AND CO-EDUCATIONAL (N = 163) SCHOOLS

| | | $\bar{X}$ | SD | t | P |
|---|---|---|---|---|---|
| (High)<br>Rated preference for English | Single<br>Co-ed | 2·41<br>2·03 | 1·12<br>1·14 | 3·73 | 0·01 |
| (High)<br>Rated preference for science | Single<br>Co-ed | 2·51<br>2·92 | 1·05<br>0·96 | −4·64 | 0·01 |
| (High)<br>Rated achievement in mathematics | Single<br>Co-ed | 3·34<br>3·50 | 0·84<br>0·92 | −2·07 | 0·05 |
| (High)<br>Rated achievement in science | Single<br>Co-ed | 3·46<br>3·27 | 0·92<br>0·82 | −2·52 | 0·05 |

For the male sample (Table 6), those from co-educational schools reported greater preference for mathematics than those from single-sex schools. Achievement in English was rated higher among males from single-sex compared to males from co-educational schools. Also, the self esteem of males was greater in single-sex compared to co-educational schools. Dependence on group scores as well as scores on the need for peer affiliation scale were higher for males from co-educational compared to single-sex schools.

Finally, in order to investigate the possible associations between locus-of-control orientations and other independent variables, correlations between these variables were calculated. Only significant correlations which were above $r = 0·30$ are considered. External locus-of-control was negatively associated with high self

TABLE 6

SIGNIFICANT DIFFERENCES BETWEEN MALES FROM SINGLE-SEX (N = 702) AND CO-EDUCATIONAL (N = 194) SCHOOLS

| | | $\bar{X}$ | SD | t | P |
|---|---|---|---|---|---|
| (High)<br>Rated preference for mathematics | Single<br>Co-ed | 2·24<br>2·48 | 1·08<br>1·02 | −2·79 | 0·01 |
| (High)<br>Rated preference for English | Single<br>Co-ed | 3·38<br>3·19 | 0·80<br>0·91 | 2·65 | 0·01 |
| (High)<br>Self esteem | Single<br>Co-ed | 2·54<br>0·10 | 3·12<br>3·22 | 3·45 | 0·01 |
| (High)<br>Peer affiliation | Single<br>Co-ed | 10·12<br>10·64 | 2·6<br>2·3 | −2·47 | 0·05 |

esteem (r = −0·52) and favourable affiliations with school (r = −0·38) and peers (r = −0·45).

No measured background variable was found to be significant in differentiating subjects in any comparisons.

## DISCUSSION

As predicted, males rated their achievement in traditional male dominated subjects (mathematics and science) higher than did females. Also in agreement with the notion that gross sex differences exist in preferences for subjects, females were found to prefer English significantly more and rate their achievement in English higher than males. However, these findings were not consistent across comparisons. Males from single-sex schools were more likely to prefer English than females from single-sex schools as well as males from co-educational schools. In addition, females from single-sex schools were more likely to prefer science and rate their achievement in mathematics and science higher than their co-educational female counterparts.

The type of school attended does seem to have differential consequences for students in terms of their stated preferences and rated achievement in subjects: those attending single-sex schools seem to be less rigidly attached to traditional views about the appropriateness of subject areas by sex. By contrast, attendance at co-educational schools appears to be associated with traditional subject preferences and related assessments of achievement in those subject areas.

The self esteem of females was found, as predicted, to be significantly lower than that of males. This difference was found to be most evident in the comparisons between males and females from single-sex schools. Contrary to expectations, no differences were found between the self esteem of females from single-sex schools and co-educational schools. However, an overall difference in self esteem between students attending co-educational and single-sex schools was apparent: those attending co-educational schools had lower self esteem than those attending single-sex schools. This result was largely due to a difference for boys.

Type of school attended for males would appear to have importance for their self esteem. Moreover, while females were found to have generally lower self esteem than males, this relationship was due to differences in single sex schools rather than co-educational. Type of school would appear to have potential as an explanatory variable in explanations of poor self esteem in adolescent females as compared with males.

The notion that females have greater needs for peer affiliation than males was not confirmed. Again, males were found to have significantly greater affiliation with peers than females for both single-sex and co-educational comparisons. As expected, males as a total group had greater affiliations with their families than females: this result was also found for single-sex school comparisons. The other comparison which proved significant in this area was that need for peer affiliation was greater for males from co-educational schools in particular.

As for self esteem, the effect of school type on needs for peer affiliation is apparent. Being male and attending a co-educational school is positively associated with high peer affiliation. By contrast, school type did not differentiate patterns of group dependence and peer affiliation for females. It would appear that explanations other than those relating to school type will need to be developed to explain this result.

Contrary to expectations, females were not found to have more external locus-of-control orientations than males. In fact, no significant differences were obtained for this variable in any comparison. The potential importance of control orientations for students' self perceptions was explored by correlating it with other dependent variables. External control orientations were negatively associated with high esteem and affiliations with family and peers. It would appear that, in some senses, externality is detrimental to perceptions of self and evaluations of school life. However, the failure for locus-of-control to differentiate students by sex or school type makes the utility of these findings questionable.

The results of the present study are limited in their generalisability to the extent that the sample was biased with respect to the socio-economic backgrounds of subjects. The consistently high social class origins of subjects need not, however, invalidate the associations reported. Indeed, despite the homogeneity on this important background variable, the complexity of relationship between school type and adolescent self perceptions remains in evidence.

For both males and females attending single-sex schools there would seem to be a greater degree of tolerance for subject areas in the curriculum which have traditionally been designated inappropriate for one sex. Moreover, attendance at single-sex schools would appear to lead to high self esteem and greater independence from peers than attendance at co-educational schools. As indicated, however, the relationship is more straightforward in the case of the male sample. The female sample did not produce consistent differences between school type and patterns of self esteem and peer affiliation.

Models relating to patterns of affiliation might offer some insight into adolescents' perceptions of school life. In particular, it would be interesting to explore further the difference in the importance that individuals attach to doing well in school (from their own and others' perspectives). In this study, females were found to perceive the value that their parents attached to doing well in school as higher than males, while males felt friends and teachers attached greater importance to their doing well than did females. These differences were replicated in the single-sex school comparisons.

In particular, the high importance that females perceived their parents attached to their doing well in school might help explain why females from single-sex schools had lower preferences and lower rated achievement in particular subjects compared to males. That is, the confidence of adolescents in their assessments of academic worth might be bound up with their dominant affiliation patterns. Self confidence in one's ability at school is more likely to be generated by those immediately involved in the context of school (peers, teachers) than by those who are not directly in a position of influence (family). Females' generally poorer self esteem, compared to males, might be related to their dominant pattern of affiliation. By contrast, males' generally favourable self esteem, relative to females, could be explained by the high importance that males perceived their teachers and peers attached to their doing well in school. This notion might also point to reasons for greater peer affiliation patterns found among males. Adolescents' perceptions of teachers', parents' and friends' attitudes would seem to be important for different subject groups and investigations in future could usefully focus on these differences as they relate to school outcome.

ACKNOWLEDGMENTS.—The author is grateful to all participants, both students and staff, for their co-operation Thanks are also extended to the National Council of Independent Schools, Canberra, Australia for financial and other assistance in the conduct of this study.

Correspondence and requests for reprints should be addressed to Dr. Anne E. Foon, School of Liberal Studies, Canberra College of Advanced Education, P.O. Box 1, Belconnen ACT 2616, Australia.

## REFERENCES

BROOM, L., and JONES, F. (1976). *Opportunity and Attainment in Australia.* Canberra: Australian National University Press.

CARPENTER, P. (1985). Single-sex schooling and girls' academic achievement *Aust NZ J. Sociol.*, 21, 456-472.

DYNAN, M. P. (1980). Do schools care? *Co-operative Research Series No. 6.* Peron· Western Australian Department of Education.

EDER, D. (1985). The cycle of popularity. *Sociol. Educ.*, 58, 154-165.

EDGAR, D. E. (1985). Adolescent competence and sexual disadvantage. *La Trobe-Sociology Papers*, Paper No. 10, La Trobe University.

FENNEMA, E., and SHERMAN, J. (1977). Sex-related differences in mathematics achievement, spatial visualization, and effective factors. *Am. educ. Res J*, 14, 51-71

GURNEY, R. (1980). Aspects of school leaver unemployment, *Melbourne Department of Psychology Paper*, Melbourne University Press.

HORNER, M. (1970). A basic inconsistency. In BARDWICK, J , DOUVAN, E., and GUTMAN, D. (Eds.), *Feminine Personality and Conflict*, California: Brooks Cole.

HORNER, M. (1972). Toward an understanding of achievement-related conflicts in women. *J. social Issues*, 28, 157-175.

MACCOBY, E., and JACKLIN, C., (1974). *The Psychology of Sex Differences.* Stanford, Calif.. Stanford University Press.

NOWICKI, S., and STRICKLAND, B. R. (1973). A locus-of-control scale for children. *J. consult. clin. Psychol.*, 40, 148-154.

PHARES, E. J. (1976). *Locus of Control and Personality*, Morristown, NJ: Silver Burdett.

PIERS, E. V. (1984). *Piers-Harris Children's Self-concept Scale.* California: Western Psychological Services

RICHMOND, P. G. (1985). The relationship of grade, sex, ability and socio-economic status to parent, peer and school affiliation, *Br. J. educ. Psychol.*, 55, 233-239.

SARAH, E., SCOTT, M., and SPENDER, D. (1980). The education of femininity: the case for single sex schools. In SPENDER, D., and SARAH, E. (Eds.), *Learning to Lose.* London: The Women's Press

WOLLEAT, P. L., PEDRO, J. D., BECKER, A D., and PENNEMA, E (1980). Sex differences in high school students' causal attributions on performance in mathematics *J. Res. Math. Educ.*, 11, 356-366.

(*Manuscript received 2nd July*, 1986)

*Br. J. educ. Psychol.*, 58, 55-68, 1988

# A PREPARATORY PROGRAMME FOR PASSAGE INTO THE ETHNICALLY INTEGRATED SCHOOL: IMPACT ON SCHOOL CLIMATE AND SOCIAL RELATIONS

By JOSEPH SCHWARZWALD, MICHAEL HOFFMAN AND TAMAR ROTEM
*(Department of Psychology, Bar-Ilan University, Israel)*

Summary. It is currently accepted that effective ethnic integration in schools requires supplemental intervention to attain its social and academic goals. The present study evaluated an Israeli preparatory programme addressing the transition into integrated junior high schools. This intensive, one-week programme of intervention prior to school entrance is aimed at improving school climate and interpersonal relationships between ethnic groups. The evaluation assessed both short- and long-term effects of both newly implemented and veteran programmes in comparison with contrast control schools. A total of 859 seventh grade students from 12 schools completed questionnaires on classroom climate and social relations, immediately following the programme and at year's end. Results indicated that the programme, especially at new sites, had continuing positive impact in reducing ethnic cleavage in interpersonal relationships and in improving perceptions of school and teacher. The benefits in perceived school climate overlaid a more general decline which appeared over time. The discussion focuses on ramifications for future intervention and evaluation with this programme model as well as its broader utility for improving school climate outside the integration context.

## INTRODUCTION

Experiences with ethnic integration in schools around the world have led to the conclusion that successful social and educational outcomes cannot be attained merely by bringing diverse students into contact in the classroom. Rather, the success of ethnic integration is dependent upon supplementary intervention to achieve basic prerequisites which foster integration aims. These include the equality of participants' status, opportunities for interpersonal intimacy, an atmosphere of co-operation, and a background of social and institutional support (Allport, 1954; Amir, 1969, 1976; Cohen and Sharan, 1980; Miller, 1980, 1981; Amir *et al.*, (1984). The present study evaluates a unique Israeli intervention programme that attempts to facilitate ethnic integration by creating an optimal transition from non-integrated neighbourhood elementary schools to integrated regional junior high schools.

Planned ethnic integration of schools in Israel came in response to specific educational problems as well as more general social tensions related to the ethnic division of the Jewish population into those from advantaged Western and disadvantaged Middle Eastern origins (Schwarzwald and Amir, 1984). At an educational level, policy-makers were faced with a persistent educational gap between Middle Eastern and Western students. A large segment of the Middle Eastern students did not attain even minimal achievement criteria set by the Ministry of Education. Moreover, an exceptionally high percentage of the Middle Eastern students dropped out voluntarily after completing the compulsory elementary grades (Smooha, 1978), It is estimated (Adler, 1984) that approximately two-thirds of the Middle Easterners dropped out of high school.

At a social level, educators and politicians were concerned over divisive patterns of intergroup prejudice and tension. These expressed themselves in reduced acceptance and even disparagement of the disadvantaged Middle Eastern groups.

Interpersonal relations were qualified by strong asymmetries. Westerners tended to prefer their own group members over Middle Easterners, whereas the Middle Easterners actually preferred Westerners over members of their own group (Schwarzwald and Yinon, 1977).

Initial efforts at altering this state of affairs took the form of programme for educational advancement and enrichment of the Middle Eastern children. However, this differential allocation of resources did not yield the desired outcome. Consequently, politicians championed a programme enforcing integration in the context of structural reform of the educational system (Blass and Amir, 1984). The two-tiered system of elementary and high schools was replaced with a three-tiered model of elementary, junior high, and high school. Integration was achieved by having the new regional junior high schools absorb students from relatively segregated neighbourhood schools located in a wide geographical spread.

This reform had two basic goals (Adler, 1984). The first was to raise achievement levels and close the educational gap by increasing the efficacy of the educational process. The second was to ameliorate tension and prejudice by providing more points of encounter between students of differing backgrounds.

Although this reform was implemented throughout the 1970s, it is yet to be completed. Furthermore, recognition has grown that structural changes permitting increased contact are not sufficient to ensure the social goals of the programme. Rather, supplementary interventions are required (Amir, 1976). Indeed, sporadic attempts have been made to adapt ethnic intervention models employed in other countries (e.g., the co-operative learning model described by Slavin, 1980, 1983). At the same time, the Israeli educational system has tried out some unique, home grown interventions (Schwarzwald and Amir, 1984; Sharan and Rich, 1984).

The one under investigation here is an intensive week-long preparatory programme aimed at facilitating the passage from the relatively segregated neighbourhood elementary school to the integrated junior high. One central aim of the preparatory programme focused on the issue of school climate (Barnholtz and Ben-Shalom, 1983). An extensive research literature has illustrated the ramifications of school climate for both academic achievement and retention as well as the tenor of interpersonal relations (McDill *et al.*, 1969; Brookover *et al.*, 1978; Epstein and Jackson, 1981; Schmuck and Schmuck, 1983). As such, improvement of school climate became an avenue for addressing the battery of goals implicit in the broader programme of educational reform.

The specific interest in school climate was dictated by recognition that Israeli schools are characterised by a traditional orientation toward academic competition (Inbar *et al.*, 1984). This climate is associated with alienation, reduced satisfaction, and lowered motivation. Furthermore, the literature indicates that such competition tends to hamper social relations across majority and minority divisions (Aronson, 1978; Slavin, 1983). In particular, the spotlight on academic status in Israeli schools puts the Middle Eastern students at a disadvantage, due to their generally weaker educational standing (Rich *et al.*, 1981).

Programme innovators were interested in altering school climate as it developed in the initial encounter between students and staff of the integrated junior high school. They desired to attain a more optimal school climate along lines much like those suggested by Schmuck and Schmuck (1983). These point to a co-operative, cohesive classroom offering students the opportunity to demonstrate their capabilities without accentuating their weakness. Additional hallmarks include an ambience of respect within a flexibly and mutually determined network of norms, as well as a framework in which all may feel themselves influential contributing members.

The eventual programme format was designed to permit students and teachers to meet one another initially in a non-educational co-operative framework which downplayed concerns of academic achievement. Further, it was devised to allow teachers and students the opportunity mutually to set flexible school and class norms in a process respecting every person's opinion and contribution. As such, the preparatory programme brought entering students together with their teachers and counsellors during the week prior to the formal start of school. A detailed series of daily activities was conducted in multiple formats, ranging from the level of dyadic encounters through class meetings and up to school-wide forums. These allowed the students to express personal interests while giving teachers an opportunity to see them in a variety of roles not often arising in the school curriculum.

The preparatory programme seemed especially worthy of evaluation because of its theoretical implications as well as its potential ease in implementation and export to other sites of ethnic integration around the world. The aim of the present study was to assess both short- and long-term effects over the school year. Here, specific attention was addressed to the programme's impact on the intertwined variables of school climate and interpersonal acceptance. One further variable of interest was the history of programme implementation in the school. Previous literature (Weiss, 1972; Anderson and Ball, 1978; Hodges and Sheehan, 1978) has suggested that veteran programmes have waning success as the initial effect of being in a pioneering experiment flags and dissipates. The evaluation addressed the influence of recent versus veteran implementation on programme outcomes.

## METHOD

### *Sample*

In 1984, the preparatory programme was operative in 30 ethnically integrated junior high schools. Eight schools were randomly selected for participation. Three were implementing the programme for the first time and the remainder had done so for at least two years. For contrast, four junior high schools were chosen from similar regions, but where no preparatory programme was operative. These schools were selected so as to ensure similarity of socioeconomic background and ethnic ratio in the classroom. Matching was based on objective data provided by the Ministry of Education and Culture.

Two seventh grade homeroom classes were randomly chosen from each school. In one school with a preparatory programme, three classes were sampled to avoid expected administrative changes in these classes. The 25 study classes encompassed 859 students. Since questionnaires were administered at two different periods, only those participating in both sessions were included in the analyses. Thus, the total number of respondents was 722. A breakdown of these by programme type, sex, ethnic origin, and academic status is given in Table 1. Ethnicity was determined on the basis of father's origin; academic status was determined by student's stream level in mathematics. Inspection of the table points out the overlap between ethnic origin and academic status in Israeli schools: Westerners tend to be placed primarily in high academic levels whereas Middle Easterners are placed primarily in lower ones.

### *Preparatory Programme*

The preparatory programme took place during the week preceding the start of the regular school year. Programme contents were based on two underlying principles — symmetry and multidimensionality — derived from the literature on informal education (Kahane, 1981) and that on the "contact hypothesis" (e.g., Allport, 1954; Amir, 1969, 1976; Miller, 1981). Symmetry refers to the equality of participant status and mutuality of goals. Symmetry is thought to enhance

TABLE 1

SAMPLE DISTRIBUTION BROKEN DOWN BY PROGRAMME TYPE, SEX, ETHNIC ORIGIN, AND ACADEMIC STATUS

| | | Programme Type | | | | | |
|---|---|---|---|---|---|---|---|
| | | Control | | New | | Veteran | |
| Ethnic Origin | Academic Status | Boy | Girl | Boy | Girl | Boy | Girl |
| Middle Eastern | Low | 25 | 24 | 8 | 15 | 20 | 21 |
| | Moderate | 26 | 25 | 15 | 13 | 35 | 29 |
| | High | 20 | 19 | 20 | 25 | 43 | 30 |
| Western | Low | 7 | 16 | 8 | 2 | 5 | 8 |
| | Moderate | 8 | 21 | 11 | 9 | 23 | 20 |
| | High | 18 | 24 | 21 | 27 | 43 | 38 |

interdependence and commitment while diverting attention from categorical sources of interpersonal diversity. Multidimensionality involves the provision for individual participation in multiple frameworks requiring varying levels of responsibility and performance. This richness of experience allows individuals to select preferred roles which present their strengths and qualities to their best advantage.

Students were asked to work in multiple frameworks: alone, as dyads or in small groups with classmates, as part of an entire class, and as part of a group of classes making up the entire grade level. Particular emphasis was placed on creating activities which would allow students to interact with classmates and teachers in multiple formats addressing the broadest possible range of activities and interests. The three overall goals were to: (a) acquaint students and staff; (b) foster class norms through direct discussion and introduction of a class contract; (c) raise class cohesion through competition with other classes. Daily activities and objectives were formally spelled out in an activity guide to ensure consistency across schools.

Activities at the level of the individual involved participation in two daily sessions of hobby and personal interest clubs. These brought together children from different classes to work and play together on themes of sport, creative work, drama, and recreation. Teachers and staff conducted and observed differing groups in order to acquaint themselves with a cross-section of children.

Activities at the small group and class level involved three daily sessions directed at observation of school functions, articulation of expectations regarding self, class, and school, and understanding differing role demands and conflicts within the school. Observation of school functions involved the preparation of dyads to follow different functionaries within the school (e.g., principal, librarian) and in its environs (bus driver, kiosk owner) to observe their role and responsibilities. These observations were reported back to the class. Articulation of expectations was carried out through informal games and small group discussions, as well as by formally writing up a class contract. Understanding role demands and conflicts was approached through role-taking games and simulations.

Activities at the level of the grade section involved both competitive and co-operative activities between whole classrooms. Competitive activities included contests in sports, athletics, music, and song. More co-operative activities focused on drama and dance. These activities were conducted in short periods at the end of each day, as well as in a class competition day held in the middle of the week.

*Measures*

*Classroom Climate Questionnaire.* The 60 items in this questionnaire were adapted from Darom and Rich's (1981) Israeli Quality of School Life Scale. Certain items and wordings were changed in order to adapt them to the programme structures surveyed.

Respondents were asked to rate each statement item on a four-point scale ranging from 1 (completely true) to 4 (completely false). Principal component analysis with varimax rotation on item scores in the initial administration yielded 10 factors (accounting for 53·7 per cent of variance). Factors included classroom satisfaction, positive teacher characteristics, classroom cohesiveness, sources of popularity, classroom companionship, school satisfaction, negative teacher characteristics, preferential treatment, antipathy to school, and learning climate. Factor items and their loadings are presented in Table 2.

TABLE 2

FACTOR ITEMS AND LOADINGS FOR CLASSROOM CLIMATE QUESTIONNAIRE

| Factor and Item Content | (% Explained Variance*) | Loadings |
|---|---|---|
| I. | Class satisfaction (20·0%) | |
| | Students in the class help each other | 0·69 |
| | Most of the students in the class like each other | 0·67 |
| | You can depend on friends in class to help you when needed | 0·59 |
| | Our class sticks together | 0·56 |
| | I have many friends in the class | 0·54 |
| | I feel good with my classmates | 0·54 |
| | I feel that I fit in well with my junior high class | 0·50 |
| | The students in the class know each other well | 0·45 |
| | In our class, even poor students are appreciated | 0·49 |
| | In our class, students like to work together | 0·47 |
| | I am satisfied with my class | 0·37 |
| | I am willing to put out a lot for the class | 0·35 |
| | In our class, a lot of new friendships have sprung up | 0·29 |
| | I wouldn't want to move to another class | 0·29 |
| II. | Positive teacher characteristics (6·6%) | |
| | Our teacher looks out for the class | 0·66 |
| | Our teacher talks things over with the class committee | 0·59 |
| | Our social classes are interesting | 0·54 |
| | I feel I can talk things over freely with my teacher | 0·51 |
| | I like my teachers | 0·49 |
| | The teachers always find time to talk with the students | 0·46 |
| | Our class has clear rules about how to behave | 0·45 |
| | The teachers have the students join in setting rules | 0·45 |
| | The teachers really listen to what students have to say | 0·42 |
| | The teachers take a personal interest in each student | 0·40 |
| | Junior high teachers are better than elementary school ones | 0·37 |
| | I wish my teachers this year could teach me next year | 0·36 |
| III. | Classroom cohesiveness (5·4%) | |
| | There are groups in our class that don't get along | 0·62 |
| | The class is broken up into cliques | 0·53 |
| | Only a minority of the students take part in class activities | 0·43 |
| | Many students are dissatisfied with the class | 0·41 |
| | There are students in class that everyone disregards | 0·39 |

TABLE 2 (Continued)

FACTOR ITEMS AND LOADINGS FOR CLASSROOM CLIMATE QUESTIONNAIRE

| Factor and Item Content | (% Explained Variance*) | Loadings |
|---|---|---|
| IV. | Sources of Popularity (3·8%) | |
| | In our class, those who get better grades are more important | 0·80 |
| | In our class, those who invest more are liked best | 0·70 |
| | The students compete with each other to be the best | 0·31 |
| V. | Preferential treatment (3·6%) | |
| | The teachers treat some of the students better than others | 0·66 |
| | The teachers treat some students worse than others | 0·52 |
| | In our class, some students are more important than others | 0·40 |
| VI. | Classroom companionship (3·0%) | |
| | I have no close friends in the class | 0·53 |
| | I spend time with class friends after school | 0·50 |
| | I have better friends in my neighbourhood than in class | 0·39 |
| VII. | School satisfaction (2·7%) | |
| | I love studying this year | 0·71 |
| | I enjoy being in school | 0·54 |
| | Many times I don't feel like going to school | 0·50 |
| VIII. | Negative teacher characteristics (2·3%) | |
| | The teachers don't trust the students | 0·51 |
| | The teachers treat us like small children | 0·46 |
| | The teachers think they're always right | 0·41 |
| | I often feel angry about my teachers | 0·39 |
| IX. | Antipathy to school (2·1%) | |
| | If I could, I would leave school | 0·74 |
| | I would be happier studying in another school | 0·70 |
| X. | Learning climate (2·0%) | |
| | In our class, the learning atmosphere is serious | 0·36 |
| | In our class, students really listen to the teacher | 0·34 |

* % Explained variance, calculated from initial factor eigenvalues.

† Nine items with loadings below 0·25 were dropped from the analysis and are not presented here

As can be seen, a number of the factors have low explained variances. This is a characteristic common to classroom climate scales (e.g., Moos, 1979; Darom and Rich, 1981), so none of the factors with notably differing contents was removed from further analysis. Nonetheless, since the items on two factors — classroom cohesiveness and classroom satisfaction — were highly similar in content, they were combined into one factor of classroom cohesiveness and satisfaction. Like considerations led to the merging of the two factors of school satisfaction and school antipathy. In both cases, merging did not alter the outcomes of the analyses. These manipulations yielded eight final factor scores in data analysis.

*Interpersonal Relationship Assessment Technique (IRAT).* The IRAT, developed by Schwarzwald and Cohen (1982), consisted of the following activities in random order. In order of increasing intimacy these include:

(a) To lend him/her a book or pen in class.
(b) To play with him/her during recess.
(c) To do homework with him/her after school.
(d) To be his/her best friend.
(e) To reveal personal secrets to him/her.

Respondents were asked whether they would or would not agree to engage in each of the IRAT activities with each same-sex classmate appearing in an alphabetic roster printed on the form.

The relative level of intimacy for each item was determined on the basis of scalogram analyses (Guttman, 1950) performed on IRAT data separately for each administration. These analyses showed that IRAT items form a unidimensional, monotonically rising scale with high coefficients for reproducibility (0·92, 0·93 for two administrations, respectively), scalability (0·71, 0·73), and reliability (0·91, 0·91).

Five separate IRAT scores were calculated for each respondent. These consisted of mean acceptance scores for Western and Middle Eastern classmates as well as means for low, moderate, and high academic status classmates.

### *Procedure*

Questionnaires were first administered to the homeroom classes two weeks after the start of school. A second administration was performed six months later. Biographical data were collected directly from school files.

## RESULTS

### *School climate*

A five-way MANOVA for repeated measures was performed on the eight factor scores for school climate. The design for this analysis was 2 (Term: short/long) × 3 (Programme: new/veteran/control) × 2 (Ethnicity: Western/Middle Eastern) × 2 (Sex: boys/girls) × 3 (Academic status: low/moderate/high). Sex and academic status were included in the analysis because of their known impact on school attitudes (Berk *et al.*, 1970; Barker Lunn, 1972).

MANOVA outcomes revealed significant main effects for programme ($F(2, 1358) = 8{\cdot}00$, $P < 0{\cdot}01$), term ($F(1, 679) = 29{\cdot}13$, $P < 0{\cdot}01$), academic status ($F(2, 1358) = 4{\cdot}56$, $P < 0{\cdot}01$), and sex ($F(1, 679) = 5{\cdot}03$, $P < 0{\cdot}01$). The single significant two-way interaction was between programme and term ($F(2, 1358) = 5{\cdot}03$, $P < 0{\cdot}01$), and the single significant triple order interaction involved programme, term, and sex ($F(2, 1358) = 2{\cdot}02$, $P < 0{\cdot}01$).

Table 3 presents mean factor scores broken down by term and programme. Inspection reveals a general decline in perceived school climate over the year. As the univariate F tests indicate (Table 4), this drop in school climate appears significantly in all factor scores other than that of Companionship.

The programme effects overlay this basic depreciation in school climate over time. Students in the preparatory programmes generally tended to perceive the school more positively than did the contrast controls, and this difference reached significance in five of the eight factors: positive teacher characteristics, negative teacher characteristics, school satisfaction, preferential treatment, and study atmosphere. These differences did not include those factors related to social and interpersonal areas.

Scheffé post hoc comparisons were conducted to assess differential impacts of new and veteran programmes. These revealed that on the five factors where effects

TABLE 3

Mean Scores for School Climate Factors for Programme by Term

| | New | | Veteran | | Control | |
|---|---|---|---|---|---|---|
| | Short | Long | Short | Long | Short | Long |
| Class Satisfaction | 1·97 | 2·22 | 2·02 | 2·26 | 2·04 | 2·17 |
| Class Companionship | 1·91 | 2·01 | 1·92 | 1·91 | 1·95 | 1·95 |
| Sources of Popularity | 2·09 | 2·02 | 2·06 | 2·22 | 1·94 | 2·21 |
| Positive Teacher Characteristics | 1·96 | 2·35 | 2·08 | 2·23 | 2·28 | 2·58 |
| Negative Teacher Characteristics | 2·09 | 2·23 | 2·16 | 2·35 | 2·27 | 2·46 |
| Differential Treatment | 2·39 | 2·55 | 2·28 | 2·78 | 2·35 | 2·75 |
| School Satisfaction | 1·71 | 2·00 | 1·78 | 2·03 | 1·92 | 2·09 |
| Learning Emphasis | 2·23 | 2·53 | 2·39 | 2·64 | 2·48 | 2·77 |

*Note:* Higher scores indicate less positive assessments.

TABLE 4

F Values in Univariate ANOVAs of School Climate Factors for Effects Found Significant in the MANOVA

| | Term | Programme | Term × Programme | Term × Programme × Sex | Term × Programme × Status |
|---|---|---|---|---|---|
| Class Satisfaction | 86·70* | 1·09 | 3·11 | < 1·00 | < 1·00 |
| Class Companionship | < 1·00 | 1·78 | < 1·00 | 1·56 | < 1·00 |
| Sources of Popularity | 20·70* | < 1·00 | 2·70 | < 1·00 | < 1·00 |
| Positive Teacher Characteristics | 167·36* | 28·48* | < 1·00 | 5·08* | < 1·00 |
| Negative Teacher Characteristics | 32·22* | < 1·00 | < 1·00 | 2·38 | < 1·00 |
| Differential Treatment | 57·30* | 3·69* | 5·12* | < 1·00 | < 1·00 |
| School Satisfaction | 64·49* | 6·06* | < 1·00 | 2·21 | < 1·00 |
| Learning Emphasis | 54·78* | 9·49* | 1·86 | 2·02 | 3·15 |

* P <0·01

were statistically significant, they were significantly greater in the new rather than veteran programme. A single exception was the case of learning emphasis.

The significant two- and three-way interactions involved variations in programme effects, but these were restricted to single factors which the factor analysis had indicated explained little variance. Nonetheless, their content differences merit examination. The three-way interaction of sex, term, and programme involved the dissipation of programme effects on boys' perceptions of positive teacher characteristics. In the beginning of the year, both boys and girls in the new programmes had significantly better evaluations of their teachers than the other students. Yet only the girls continued to do so at year's end.

The two-way interaction of programme by term reflects a sleeper effect in the single factor of preferential treatment. Although neither new nor veteran programmes had a short-term effect on this factor, a significant long-term effect arose in the new programmes at year's end. Students in the new programme did not display as great a decline in this perception as did the students in the veteran programmes or the contrast controls.

The MANOVA indicated significant main effects for sex and academic status that were not directly related to the study proper, yet warrant consideration because of their implications for the study's robustness and integrity. The sex effects consisted of a generally more negative perception of school climate among boys than girls. This difference reached statistical significance for three of the eight factors: sources of popularity (M boys = 2·17 *vs.* M girls + 1·95), negative teacher characteristics M = 2·27 *vs.* 2·08), and school satisfaction (M = 1·87 *vs.* 1·76). These findings parallel those reported by Barker Lunn (1972), Berk *et al.* (1970), and Rich and Darom (1981).

Academic status effects were restricted to the two factors of positive teacher characteristics and class companionship. In regard to teacher characteristics, low-status students held more positive estimations than did either moderate- or high-status students (M = 2·00, 2·18, and 2·15, respectively). An inverse pattern arose in regard to classroom companionship: low-status students expressed less satisfaction with class relations than did moderate- or high-status students (M = 2·15, 1·96, and 1·80, respectively).

### *Social acceptance*

The analysis of social acceptance focused on the question of whether participation in a preparatory programme ameliorated ethnic biases in interpersonal acceptance. Since the academic status of one's classmate is known to moderate the degree of interethnic acceptance (Schwarzwald and Cohen, 1982; Hoffman and Schwarzwald, in press), this variable was also included in the analysis. Thus, the initial ANOVA design of IRAT scores involved six variables of programme type, term, target origin, target academic status, respondent origin, and respondent sex. Since this analysis indicated significant sex effects, separate five-way ANOVAs were performed on boys' and girls' scores. In these, low and moderate academic status target students were combined into one category since the overlap of ethnicity and academic status often yielded very low frequency cells for low status Western targets (see Table 1). Because of incomplete IRAT questionnaires, the number of subjects in the ANOVAs dropped to 598. Significant effects in the ANOVAs for boys and girls are summarised in Table 5.

TABLE 5

SIGNIFICANT F VALUES (P < 0·01) IN ANOVAS OF IRAT SCORES

| Effect | Boys | Girls |
|---|---|---|
| Target origin | 10·61 | 21·12 |
| Target academic status | 99·39 | 214·70 |
| Programme × target origin | 8·04 | 5·35 |
| Programme × target academic status | NS | 4·64 |
| Respondent origin × target academic status | 7·41 | NS |

*Note:* df error term = 253 for boys and 313 for girls.

The significant main effect of target origin is subsumed in the interaction of programme type and target origin. As can be seen from the relevant means in Table 6, there was a tendency in the control groups for boys and girls to show the often reported (Schwarzwald and Amir, 1984) preference for Western over Middle Eastern classmates. Among boys, this preference was found in both new and veteran preparatory programmes. For girls, however, only the new programme mitigated ethnic preference; girls in the veteran programmes displayed a level of categorical preference equal to that in the control groups.

TABLE 6

MEAN IRAT SCORES FOR PROGRAMME TYPE AND TARGET ORIGIN, BY RESPONDENT SEX

| | Programme Type and Target Origin | | | | | |
|---|---|---|---|---|---|---|
| | Control | | New | | Veteran | |
| Respondent Sex | W | M-E | W | M-E | W | M-E |
| Boys | 3·34 | 3·02 | 3·10 | 3·02 | 3·14 | 3·13 |
| Girls | 3·55 | 3·39 | 3·43 | 3·42 | 3·46 | 3·17 |

The dramatic influence of target's academic status reflects a preference for high over low academic targets among both boys and girls. The preparatory programmes did not appear to reduce this preference. Indeed, as the means in Table 7 suggest, this preference was even more marked among girls from veteran preparatory programmes than in the control group. However, this interactive effect of programme type and target status was significant only for girls.

TABLE 7

MEAN IRAT SCORES FOR PROGRAMME TYPE AND TARGET ACADEMIC STATUS, BY RESPONDENT SEX

| | | Programme Type | | |
|---|---|---|---|---|
| Respondent Sex | Target Academic Status | Control | New | Veteran |
| Boys | High | 3·41 | 3·20 | 3·38 |
| | Low | 2·95 | 2·92 | 2·90 |
| Girls | High | 3·74 | 3·67 | 3·70 |
| | Low | 3·19 | 3·18 | 2·92 |

The analysis for boys yielded one final interaction between respondent origin and target academic status. The preference for high over low academic classmates was more marked among Western respondents (M = 3·46 *vs.* 2·91 for high *vs.* low status targets) than among Middle Eastern ones (M = 3·25 *vs.* 2·94). This interaction did not reappear among girls.

TABLE 8

MEAN IRAT SCORES BY PROGRAMME TYPE, RESPONDENT AND TARGET ACADEMIC STATUS

| Respondent Status | Target Status | Programme Type<br>New | <br>Veteran | <br>Control |
|---|---|---|---|---|
| | Low | 3·27 | 3·12 | 3·12 |
| Low | Moderate | 3·21 | 3·04 | 3·26 |
| | High | 3·16 | 3·31 | 3·56 |
| | Low | 2·26 | 2·94 | 3·13 |
| Moderate | Moderate | 3·58 | 3·05 | 3·34 |
| | High | 3·70 | 3·47 | 3·60 |
| | Low | 2·48 | 2·76 | 2·87 |
| High | Moderate | 2·89 | 2·93 | 3·22 |
| | High | 3·58 | 3·71 | 3·44 |

## DISCUSSION

Review of the study findings indicates beneficial impacts in both school climate and interethnic relationships as a result of the short, but intensive intervention in the passage from ethnically homogeneous to ethnically integrated schools. These gains were particularly marked among schools employing the preparatory programme for the first time. Further, the changes in attitudes were most noted among boys rather than girls.

In assessing these patterns, it must be recalled that the central intervention objective was to move the school climate from an orientation of competition and authority to one of co-operation and mutual respect. This objective was a goal in its own right as well as means for improving interethnic social relations. The present results indicate that this goal was attained, at least in respect of the children's perception of teachers and the school. Children who underwent the preparatory programme held more positive attitudes toward their teachers and their relationship with them, as well as toward the school's learning environment.

These gains in school climate take on added import in light of the general decline in perceptions of the school over the school year. The students in this study tended to be less positive about all aspects of school life as the school year progressed. Even though the preparatory programme did not sway this decline, its impact continued to be felt throughout the year.

The decline in perceived school climate with time has been reported in other works (Epstein and Jackson, 1981). Sharan and Hertz-Lazarowitz (1981) have argued that it is due to basic structural factors in pedagogic approach which sully the quality of relationships and impair satisfaction. As such, it is not surprising that the preparatory programme did not allay the drop over time, as it did not involve any alteration of curricula or instructional methods in the course of the regular school year. Nevertheless, it is impressive that the initial gains of a short, one-week intervention were not engulfed or eradicated in the face of these long-standing, contrary forces.

As expected, gains in school climate were accompanied by a reduction of ethnic bias in interpersonal acceptance. Although Western and Middle Eastern students in the control groups displayed the frequently reported preference for their Western classmates (Schwarzwald and Amir, 1984), this pattern was dissipated among students participating in the preparatory programmes. As in the case of school climate effects, the changes in interpersonal relationships figured more prominently in the newly implemented preparatory programmes rather than the more veteran ones. Further, gains were more noted among boys than girls.

Despite the reduction of ethnic bias, it is interesting to recall that the preparatory programmes did little to change the influence of academic status on social relations. All students tended to prefer classmates from higher academic streams, regardless of ethnic origin. As such, it appears that the preparatory programmes did not necessarily blind students to achievement considerations. Rather, as expected, they moved the basis for preference from considerations involving biases of ascribed group features to ones resting on achieved individual qualities.

In this regard, it must be noted that in Israel, as in other countries, there is a major overlap between ethnicity and academic achievement. This overlap may tend to mask programme effects on ethnic bias from both the naïve observer as well as the more sophisticated researcher. Only by carefully controlling for issues of academic status can one gauge the true impact of ethnic interventions in school settings (cf. Hoffman and Schwarzwald, in press).

The finding that the programme had greater impact on boys than girls deserves remark. One possibility is that the programme's content was more attractive to boys than girls. Many of the activities involved group sport or athletics. These may have been less inviting to girls and offered a less potent base for direct contact and interaction for them.

The particular efficacy of the newly implemented programmes also warrants consideration. In almost all cases where the preparatory programme led to changes, these were more marked in the new rather than veteran programmes. Similar findings on the heightened effects of programmes in their initial stages have appeared in evaluations of other educational and social action programmes (Weiss, 1972; Anderson and Ball, 1978; Hodges and Sheehan, 1978). Common explanations focus on issues of greater motivation associated with new and experimental efforts. Discussion with programme staff also suggests the importance of ongoing in-house training and supports surrounding programme installation. The veteran staff members no longer benefited from supervision and were less aware of programme improvements that had been introduced with time.

This decline in programme efficacy with programme seniority has important practical implications. From an evaluation perspective, it indicates that evaluation of new site implementations may lead to an overestimate of programme efficacy over the long run. From a programming viewpoint, it suggests the possible need for "refresher" meetings with veteran staff even after an initial history of implementation success.

On the whole, it appears that the preparatory programme holds promise. Its benefits in terms of improved school climate should not be underestimated, particularly in light of the major value attributed this variable by educators and psychologists. Although the aegis for this programme was in regard to social integration, future contexts for implementation are by no means limited to this frame of reference. Indeed, the low cost and simplicity of this programme make it a good candidate for export to other frameworks concerned with deteriorating school climate.

ACKNOWLEDGMENT.—The research was supported by a grant from the Israel Ministry of Education and Culture.

Correspondence and requests for reprints should be addressed to Professor Joseph Schwarzwald, Department of Psychology, Bar-Ilan University, 52 100 Ramat Gan, Israel.

## REFERENCES

ADLER, C. (1984). School integration in the context of the development of the Israeli educational system. In AMIR, Y., SHARAN, S., and BEN-ARI, R. (Eds.), *School Desegregation: Cross-cultural Perspectives*. Hillsdale, NJ: Erlbaum.

ALLPORT, G. W. (1954). *The Nature of Prejudice*. Cambridge, Mass. Addison-Wesley.

AMIR, Y. (1969). Contact hypothesis in ethnic relations. *Psychol. Bull.* 71, 319-342.

AMIR, Y. (1976). The role of intergroup contact in change of prejudice and ethnic relations In KATZ, P. A. (Ed.), *Towards the Elimination of Racism*. New York: Pergamon Press.

AMIR, Y., SHARAN, S., and BEN-ARI, R. (1984). Why integration? In AMIR, Y., SHARAN, S., and BEN-ARI, R. (Eds.), *School Desegregation: Cross-cultural Perspectives*, Hillsdale, NJ: Erlbaum.

ANDERSON, S. B., and BALL, J (1978). *The Profession and Practice of Programme Evaluation*. San Francisco: Jossey-Bass.

ARONSON, E. (1978). *The Jigsaw Classroom*. Beverly Hills, Calif.: Sage.

BARKER LUNN, J. (1972). The influence of sex, achievement level, and social class on junior high school children's attitudes. *Br. J. educ. Psychol.*, 42, 70-74.

BARNHOLTZ, Y., and BEN-SHALOM, G. (1983). Preparatory programme for social integration in the seventh grade of junior high school. Jerusalem: Ministry of Education and Culture. (Hebrew).

BERK, L. E., ROSE, H. M., and STEWART, D. (1970). Attitudes of English and American children toward school experience. *J. educ Psychol.*, 61, 33-40.

BERNDT, T. J. (1982). The features and effects of friendship in early adolescence. *Child Dev.*, 53, 1447-1460.

BLASS, N. J., and AMIR, B. (1984). Integration in education: the development of a policy. In AMIR, Y., SHARAN, S., and BEN-ARI, R. (Eds.), *School Desegregation: Cross-cultural Perspectives*. Hillsdale, NJ: Erlbaum.

BRAND, E. S., RUIZ, R A., and PADILLA, M. A. (1974). Ethnic identification and preference: a review. *Psychol. Bull.*, 81, 860-890.

BROOKOVER, W., SCHWEITZER, H., SCHNEIDER, M., and BEADY, E. (1978). Elementary school social climate and school achievement. *Am educ. Res. J.*, 15, 301-318.

COHEN, E. G., and SHARAN, S. (1980). Modifying status relations in Israeli youth: an application of expectation states theory. *J. cross-cult. Psychol.*, 11, 364-384.

DAROM, E., and RICH, Y. (1981). Development and validation of the Israeli Quality of School Life Scale. In EPSTEIN, J. L. and JACKSON, P. W. (Eds.), *The Quality of School Life* Lexington, Mass.· Heath.

EPSTEIN, J. L , and JACKSON, P. W. (Eds.), (1981). *The Quality of School Life*. Lexington, Mass.: Heath.

GUTTMAN, L. (1950). The basis for scalogram analysis. In STOFER, S A. (Ed.), *Measurement and Prediction* Princeton, NJ: Princeton University Press.

HODGES, W. L., and SHEEHAN, R. (1978). Evaluation: strategies for generating knowledge. *New Directions for Program Evaluation*, 3, 81-93.

HOFFMAN, M., and SCHWARZWALD, J (in press). Moderating effects of educational standing on interethnic relations. *Int. J. Intercultural Relations*.

INBAR, D., RESH, N., and ADLER, C. (1984). In AMIR, Y., SHARAN, S., and BEN-ARI, R. (Eds.), *School Desegregation: Cross-cultural Perspectives*. Hillsdale, NJ: Erlbaum.

KAGAN, S. (1977). Social motives and behaviors of Mexican-American and Anglo-American children. In MARTINEZ, J. L. (Ed.), *Chicano Psychology*. New York: Academic Press.

KAHANE, R. (1981). Informal youth organizations: a general model. *Sociol. Inquiry*, 4, 17-28.

McDILL, E. L., MEYERS, E. D., and RIGSBY, L. C. (1969). Educational climates of high schools: their effects and sources. *Am J. Sociol.*, 74, 567-586.

MILLER, N. (1980). Making desegregation work. In STEPHAN, W., and FEAGIN, J. (Eds.), *School Desegregation: Past, Present and Future*. New York: Plenum.

MILLER, N. (1981). Changing views about the effects of school desegregation: brown then and now. In BREWER, M. B., and COLLINS, B. E. (Eds.), *Scientific Inquiry and the Social Sciences*. San Francisco: Jossey-Bass.

MOOS, R. H. (1979). *Evaluating Educational Environments*. San Francisco: Jossey-Bass.

RICH, Y., AMIR, Y., and BEN-ARI, R. (1981). Social and emotional problems associated with integration in Israel junior high schools. *Int. J. Intercultural Relations*, 5, 259-275.

RICH, Y., and DAROM, E. (1981). Pupils' perception of the quality of life in advantaged and disadvantaged Israeli schools. In EPSTEIN, J. L., and JACKSON, P. W. (Eds.), *The Quality of School Life*. Lexington, Mass.: Heath.

SCHMUCK, R. A., and SCHMUCK, P. A. (1983). *Group Processes in the Classroom*. Dubque, Iowa: Brown Company.

SCHWARZWALD, J. (1984). Integration as a situational contingent: secular versus public education. In AMIR, Y., SHARAN, S., and BEN-ARI, R. (Eds.), *School Desegregation: Cross-cultural Perspectives*. Hillsdale, NJ: Erlbaum.

SCHWARZWALD, J., and AMIR, Y. (1984). Interethnic relations and education: an Israeli perspective. In MILLER, N., and BREWER, M. B. (Eds.), *Groups in Contact: The Psychology of Desegregation*. Orlando, Flor.: Academic Press.

SCHWARZWALD, J., and COHEN, S. (1982). Relationship between academic tracking and the degree of interethnic acceptance. *J. educ. Psychol.*, 74, 588-597.

SCHWARZWALD, J., and ROTEM, T. (1985). *The Impact of School Social Preparatory Intervention Programme on School Climate, Social Relationships, Self-concept and Educational Aspirations*. Jerusalem: Ministry of Education and Culture/Department of Youth.

SCHWARZWALD, J., and YINON, Y. (1977). Symmetrical and asymmetrical interethnic perception in Israel. *Int. J. Intercultural Relations*, 1, 40-47.

SHARAN, S., and HERTZ-LAZAROWITZ, R. (1981). *Changing Schools*. Tel-Aviv: Ramot. (Hebrew).

SHARAN, S., and RICH, Y. (1984). Field experiments in ethnic integration in Israeli schools. In AMIR, Y., SHARAN, S., and BEN-ARI, R. (Eds.), *School Desegregation: Cross-cultural Perspectives*. Hillsdale, NJ: Erlbaum.

SLAVIN, R. E. (1980). Co-operative learning. *Rev. educ. Res.*, 50, 315-342.

SLAVIN, R. E. (1983). *Co-operative Learning*. New York: Longman.

SMOOHA, S. (1978). *Israel: Pluralism and Conflict*. London: Routledge and Kegan Paul.

WEISS, C. H. (1972). *Evaluation Research*. Englewood Cliffs, NJ: Prentice Hall.

ZIEGLER, S. (1982). The effectiveness of co-operative learning teams for increasing cross-ethnic friendship. *Hum. Organization*, 40, 264-268.

(*Manuscript received 2nd March*, 1987)

*Br. J. educ. Psychol.*, 58, 69-77, 1988

# THE RELATIVE SEVERITY OF COMMON CLASSROOM MANAGEMENT STRATEGIES: THE STUDENT'S PERSPECTIVE

By MOSHE ZEIDNER
*(School of Education, University of Haifa, Israel)*

Summary. The major aim of this study was to assess students' perceptions of the severity of various classroom management strategies frequently employed by teachers in dealing with student misbehaviours, and to gauge the specific dimensions underlying student perceptions. The sample consisted of 281 junior high school students in Israel who rated the severity of each of 24 commonly found classroom management strategies via a 5-point Likert-type rating scale. In addition, for comparative purposes, ratings of the management strategies were obtained for a sample of 80 junior high school teachers and their mean ratings compared with those of the student sample. A principal factor analysis of the scale ratings revealed five meaningful dimensions underlying student perceptions: (a) student referral to school or parental authority; (b) punitive teacher reaction; (c) assigning students unpleasant classroom related tasks; (d) placing constraint on students' time; and (e) assigning students unpleasant extra-class assignments. Mean ratings of the various management strategies were relatively unaffected by the students' demographic characteristics. Similarly, the rank order of students' mean ratings was remarkably invariant across sex and social group membership. Whereas students' and teachers' mean ratings were significantly discernible for about half the items on the inventory, the two groups rank ordered the classroom management techniques quite similarly with respect to their perceived severity, implying similar classroom management strategy hierarchies for the two groups. The implications of the findings for classroom control are discussed and explicated.

Classroom discipline problems appear to have plagued school teachers and administrators from time immemorial and will most likely continue to do so in the near future with unrelenting severity. Again and again, teachers have been reported to judge student misbehaviour and classroom discipline to be among the most difficult and disturbing aspects of the teaching experience (Clarizio, 1976) as well as a major factor contributing to teacher discontent and burnout (Wolfgang and Glickman, 1980). Indeed, adequate classroom discipline is generally regarded to be one of the most essential aspects of education (Clarizio, 1976; Dreikurs *et al.*, 1982) as well as an absolute prerequisite to achieving instructional objectives and safeguarding students' psychological, social and physical well-being (Clarizio, 1976). As experts have pointed out (Clarizio, 1976) the relatively rudimentary stage of the science of classroom discipline may account for the often inadequate preparation of teachers in classroom control techniques and the consequent difficulties in implementing these strategies in the classroom.

It appears that a variety of factors combine to influence and shape teachers' observed classroom management techniques and behaviours, including: (a) formally acquired strategies, generally based on learning theory and psychodynamic principles; (b) teacher tradition and lore; and (c) intuition and personal experience (Reilly and Lewis, 1983). As a rule, the views and perceptions about classroom discipline techniques held by the prime targets and recipients of teacher sanctions, i.e., the school students themselves, are hardly a major factor influencing teacher management behaviour. As Dreikurs *et al.* (1982) have pointed out, it is truly paradoxical that the very students disciplined are completely left out of the picture, with teachers confronting the student with consequences decided upon by the teacher, regardless of whether or not they make sense to the student. In view of this, it follows that students, who are the recipients of classroom discipline, ought to have their views thoroughly explored to establish how they perceive and react to various commonly employed discipline techniques in the classroom. Furthermore, it is assumed that teachers ought to take the students' perspective into consideration, when meting out punishments or when deciding upon a particular control technique as appropriate for classroom discipline. This appears to be particularly true in view of the currently fashionable "interactionist" phase in classroom inquiry, stressing phenomenological and ethno-methodological perspectives, coupled with the assumption that students' general school attitudes and dispositions are influenced by disciplinary strategies used by teachers. Unfortunately, at present, school personnel do not appear to be sufficiently aware of the students' perspective on classroom discipline and it is seldom used as a guideline for choosing appropriate classroom management techniques.

Indeed, if teachers and students differ meaningfully in their perceptions of classroom management along a variety of dimensions (e.g., perceived severity, perceived effectiveness, perceived dislike, etc.), the sanctions meted out by teachers in response to specific student misdemeanours may not always make sense to the student and may be viewed as more (or less) severe than intended. If students fail to agree with the teacher's disciplinary action on account of varying perceptions of the control strategy employed by the teacher, the teacher's sanctions may miss the mark. Indeed, incongruence between teacher and student perceptions of various control techniques may lead students to feel that the punishment meted out by the teacher is unfair or unjustified, thereby resulting in possible student anger, discontent, and further student defiance and classroom misbehaviour.

Moreover, teachers may take for granted that the classroom management control hierarchy they have in mind, which partly determines their disciplinary actions, is similarly meaningful to students of varying sex and social class background. However, the question looms large whether or not classroom strategies are similarly perceived by students as a function of sex or social group membership, particularly in view of the reported differential socialisation practices and experiences of various social and gender groups (Anastasi, 1958).

Whereas considerable attention has been accorded in the literature to perceptions of student misbehaviours and to the degree to which various parties (e.g., teachers, students, clinicians) hold different attitudes towards students' behaviour problems (Beilen, 1959; Ziv and Spigel, 1969), relatively little comparable attention has focused on the other side of the coin, namely — student and teacher perceptions of classroom management techniques. Aside from a sprinkling of studies (Burns, 1978; O'Hagan and Edmunds, 1982), employing relatively simple methods of data collection and analysis, little empirical research has focused on students' perception of common classroom management techniques or their demographic correlates. In addition, very little is known about the sex or social class correlates of students' perceptions or about the specific dimensions underlying students' perceptions. As McCann (1978) commented, following a review of the related literature, "we are left with a most significant research area that deserves more attention that it has yet received" (p. 171).

In view of the gaps in the literature, the primary aim of this study was to investigate classroom management techniques from the students' perspective. Specifically, the following problems will be addressed:

(1) How do students perceive commonly used classroom management techniques with respect to their relative degree of severity?

(2) Are there meaningful sex and social-class differences in the perceptions of classroom management techniques?

(3) Do students and teachers differ in their ratings of common classroom management techniques with respect to their perceived severity?

(4) What are the specific dimensions underlying students' judgments of the severity of classroom management techniques?

## METHOD

### *Sample*

The student sample consisted of 281 junior high school students drawn from four community schools in Northern Israel. Three of the schools cater primarily to middle-class populations of European cultural background, whereas the fourth school caters mainly to lower-class populations of Asian and African extraction. Accordingly, the majority of the student sample was of middle-class background (69 per cent, N = 194), whereas the remainder was of lower-class family background (31 per cent, N = 87), as determined by joint consideration of the educational level and occupational status of the major breadwinner in the student's family. The sample was about equally divided by sex (males, 46 per cent, females, 54 per cent) but unequally divided by grade level (7th, 54 per cent; 8th, 20 per cent; 9th, 26 per cent).

### *Questionnaire construction and administration procedures*

A classroom management inventory, specifically constructed and pretested for the purpose of this study, served as the main instrument in assessing the perceived severity of classroom management techniques. An exploratory study was first conducted in order to help identify some of the more salient and frequently used management techniques prevalent in the Israeli junior high school classroom. To that end, 60 students from three comprehensive junior high schools in Northern Israel were interviewed and were asked to report in detail incidences of disruptive classroom behaviour they witnessed or were directly involved in over the past six months, which resulted in a disciplinary action response on the part of the classroom teacher on duty. Disruptive behaviour was defined to the students as "any student behaviour which seriously interfered with the teaching process or upset the normal running and management of the classroom". The interviews were conducted by six student teachers, each interviewing 10 students randomly chosen from their respective homeroom classes. The six classes chosen were deemed by the student-teacher training supervisor to be representative of junior high school classes in the school district in which this study was conducted.

The following forms of classroom misbehaviour were most frequently mentioned (each by at least 5 per cent of the sample), as resulting in disciplinary action on the part of the teacher: (a) excessive talking out of turn; (b) non-completion of homework assignments; (c) forgetting school material; (d) littering school ground; (e) damaging or destroying school property; (f) passing out notes in class; (g) talking back to teacher; (h) dishonesty or lying to teacher; (i) physical aggression towards peers; (j) verbal aggression towards peers; (k) non-completion of school tasks; (l) breaking school social codes or violating school norms; (m) indecent language or gestures; (n) tardiness to class or school activities; (o) absenteeism from class; (p) inattention to classroom activity; (q) throwing objects; (r) eating, drinking, chewing gum, or smoking in class; (s) vandalising school property; (t) violating school regulations; (u) stealing school property; (v) cheating in examinations; (w) pilfering from classmate; (x) forgery of signature (or documents).

On the basis of a detailed analysis of the interview data detailing reported teachers' responses to classroom misbehaviour, the following classroom management strategies were the most frequently obtained and therefore considered for inclusion in the inventory; (a) permanent suspension from school; (b) shaming the student; (c) permanent removal from class; (d) parent-principal conference; (e) temporary demotion to lower grade; (f) denial of special privileges; (g) communicating misbehaviour to parents; (h) lowering mark; (i) evicting student from class; (j) learning material by heart; (k) summoning student to principal's office; (l) surprise quiz; (m) busy work; (n) detention after class; (o) teacher-student conference; (p) additional monitor duty; (q) cleaning up school grounds; (r) shortening recess; (s) reporting to school at an early hour; (t) non-verbal communication from teacher; (u) verbal reprimand; (v) additional homework; (w) being placed in the corner; (x) special classroom seating.

Students were presented with a list of the 24 most commonly occurring classroom management strategies detailed above and were requested to rate each strategy with respect to "perceived severity" along a 5-point Likert type scale (5 = very severe, 1 = not at all severe). In addition, students were asked to fill out a short personal data inventory aimed at gathering relevant demographic information (i.e., sex, parental education and occupation, grade, level, etc.). Also, students were asked to evaluate themselves with respect to academic achievement in school along a 5-point scale (5 = way above average; 1 = way below average). The draft inventory was presented to 10 junior high school teachers for comment and then pretested on a sample of 20 students for clarity of instructions and comprehensibility.

The final version of the inventory was group-administered to the final sample of 281 students during regular homeroom class period by teachers-in-training assigned to the respective classrooms.

In addition, for comparative purposes, the same classroom management inventory was administered to a total of 80 female junior high school teachers teaching in the same school district from which the student sample was drawn.

### *Statistical analysis*

Group differences on individual item ratings were tested via t-tests for independent samples. The factors underlying management strategy severity ratings were determined via principal factor analysis with varimax rotations.

## RESULTS

### *Student ratings of classroom strategies*

Table 1 presents the mean ratings and SDs obtained for the set of 24 management strategies rated by students and teachers with respect to severity. The item means for student ratings ranged from a high of M = 4·8 (permanent eviction from school) to a low of M = 1·7 (change in student's seating arrangement).

The management strategies lend themselves to classifications into three main categories, on the basis of the students' median severity ratings.

(1) The first category encompasses strategies students view as "very severe" or "severe" (i.e., accorded median ratings of 5 or 4), such as: permanent student suspension from school, personal insult of student, permanent student transfer to different class, summoning parents to confer with principal, temporary demotion to lower class, removal of privileges, bringing problem to the attention of parents, lowering school marks and temporary eviction from class.

Interestingly, both male and female students as well as students of varying

TABLE 1

Mean Ratings of Common Classroom Management Strategies

| | Strategies | Students (N = 281) | | Teachers (N = 80) | | | |
|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | d[a] | t[b] |
| 1. | Permanent suspension from school | 4·78 | 0·59 | 4·60 | 0·71 | 0·28 | 2·29** |
| 2. | Shaming or personally insulting student | 4·32 | 0·94 | 3·14 | 1·04 | 1·19 | 9·66*** |
| 3. | Permanent removal from class | 4·30 | 1·03 | 4·18 | 0·98 | 0·12 | 0·93 |
| 4. | Parent-principal conference | 4·21 | 0·95 | 3·14 | 0·94 | 1·13 | 8·09*** |
| 5. | Temporary demotion to lower grade | 4·03 | 1·20 | 3·91 | 0·94 | 0·11 | 1·79 |
| 6. | Withdrawal or denial of special privileges | 4·00 | 1·05 | 3·74 | 1·21 | 0·23 | 1·88 |
| 7. | Communicating problem to parents | 3·46 | 1·05 | 3·34 | 1·08 | 0·11 | 0·89 |
| 8. | Lowering school mark | 3·46 | 1·05 | 4·03 | 1·05 | −0·54 | −4·28** |
| 9. | Throwing student out of class | 3·40 | 1·14 | 3·65 | 1·09 | −0·22 | 1·75 |
| 10. | Learning material by heart | 3·22 | 1·12 | 2·90 | 1·02 | 0·30 | 4·88*** |
| 11. | Summons of student to principal's office | 3·17 | 1·32 | 2·80 | 0·90 | 0·33 | 2·35* |
| 12. | Surprise quiz | 3·09 | 1·21 | 3·06 | 1·14 | 0·03 | 0·20 |
| 13. | Busy work | 2·93 | 1·13 | 2·63 | 1·02 | 0·28 | 2·14 |
| 14. | Detention after class | 2·90 | 1·00 | 3·00 | 1·10 | −0·10 | 0·77 |
| 15 | Teacher-student conference | 2·70 | 1·11 | 3·08 | 1·03 | 0·36 | −2·74*** |
| 16. | Additional monitor duty | 2·70 | 1·11 | 2·66 | 1·08 | 0·04 | 0·29 |
| 17. | Cleaning up school grounds | 2·60 | 1·11 | 2·59 | 1·07 | 0·01 | 0·67 |
| 18. | Shortening recess | 2·59 | 1·23 | 2·41 | 1·02 | 0·16 | 1·20 |
| 19. | Reporting early to school | 2·52 | 0·91 | 2·55 | 0·98 | −0·03 | −0·26 |
| 20. | Teacher's aversive non-verbal communication | 2·49 | 1·15 | 2·93 | 1·04 | −0·40 | −3·08* |
| 21. | Verbal reprimand | 2·47 | 1·11 | 2·88 | 1·04 | −0·38 | −2·95** |
| 22. | Additional homework | 2·22 | 0·93 | 2·22 | 1·05 | 0·00 | 0·00 |
| 23. | Being placed in corner | 1·90 | 0·95 | 2·70 | 0·97 | −0·83 | −6·61** |
| 24. | Special classroom seating | 1·73 | 0·92 | 2·16 | 0·95 | −0·46 | −3·66*** |

*P < 0·05 **P < 0·01 ***P < 0·001

*Notes:*

a. "d" designates the mean difference between student and teacher ratings (student-teacher) divided by the within-group standard deviation.

b. t-test for independent groups conducted with 359 d.f.

social group membership view "permanent suspension from school" and "personal insult to student" as the two most severe sanctions appearing on the inventory (Mdn = 5), with 95 per cent and 85 per cent of the sample, respectively, judging these items as "very severe".

(2) The second category of classroom strategies contains items viewed to be

"moderately severe" (i.e., accorded a median rating of 3), such as: additional student monitor duty, reporting to class early, after class detainment, assignment of busy work, threatening teacher glares, student-teacher conference or reporting to the principal's office. The common theme uniting many of the strategies above was placing "constraints" on students' degrees of freedom and time, either within the school setting (e.g., coming in early to school, staying after school) or outside of school (e.g., busy work).

(3) The third category includes teacher management strategies viewed as relatively unsevere (i.e., accorded median ratings of 2 or 1) such as: rearranging student's seating position, placing student in a corner, assigning student extra homework, shortening recess, assigning student to clean school grounds, or verbally reprimanding students. Thus, control techniques involving the removal of privileges or assignment of aversive task and verbal punishment were viewed as relatively low in terms of severity.

Hotelling's $T^2$ was used to test for differences in mean scale rating vectors by sex and social class. The analysis showed non-significant effects for sex, social class, and their interaction. Therefore ratings of the various management techniques considered together are not meaningfully different for varying sex or social subgroups. Furthermore, separate analysis of individual items by sex and social class revealed few meaningful effects. Also, perceived academic status in school did not correlate significantly with any of the rating scales.

The rank order correlations between the mean ratings of the various management techniques for males and females ($\rho = 0 \cdot 97$), on one hand, and for middle-class and lower-class students ($\rho = 0 \cdot 96$), on the other, indicate that background variables have little effect on the rank order of the mean ratings.

An inspection of Table 1 shows significant differences between student and teacher mean ratings for about half the items on the inventory. However, students and teachers rank ordered the classroom management techniques quite similarly ($\rho = 0 \cdot 84$), implying comparable perceptions of classroom management hierarchies for the two groups. On the one hand, students compared to teachers rated the following techniques meaningfully higher, with respect to severity: (a) shaming the student ($d = 1 \cdot 19$) and parent-principal conference ($d = 1 \cdot 13$). On the other hand, teachers accorded meaningfully higher severity ratings than students to the following techniques: (a) being placed in the corner ($d = -0 \cdot 83$) and lowering school marks ($d = -0 \cdot 54$).

### *Underlying structure of student ratings*

In order to uncover empirically the specific dimensions underlying student perceptions of various management strategies with respect to severity, the ratings of the 24 management strategies were subjected to a principal axis factor analysis. An eigenvalue of 1 served as the smallest admissible value for the retention of an extracted factor in the analysis, with factor loadings of 0·35 considered minimal loadings for inclusion of the item in factor interpretation.

As shown in Table 2, the following five principal factors were extracted:

(1) The first labelled "Conferring with Authority" was marked by three main management strategies, namely: parent-principal conference, parent-teacher conference and student-teacher conference, with factor loadings of 0·84, 0·68 and 0·61 respectively. In each of the said strategies, the discipline problem was handled by way of verbal communication or dialogue between the student and his parents or a school authority representative (i.e., teacher or principal).

(2) The second factor derived, labelled "Aversive Teacher Response", involves

TABLE 2

PRINCIPAL FACTOR LOADINGS OF STUDENTS' MANAGEMENT STRATEGY RATINGS

| | Strategy | F1 | F2 | F3 | F4 | F4 | $h^2$ |
|---|---|---|---|---|---|---|---|
| 1. | Permanent suspension | 0·29 | 0·13 | 0·09 | 0·18 | 0·11 | 0·16 |
| 2. | Personal insult | −0·11 | 0·23 | 0·33 | 0·21 | −0·07 | 0·22 |
| 3. | Eviction from class | 0·26 | 0·03 | 0·21 | 0·11 | 0·00 | 0·12 |
| 4. | Parent-principal conference | 0·84 | 0·01 | 0·06 | 0·03 | 0·06 | 0·72 |
| 5 | Demotion to lower grade | 0·13 | 0·03 | 0·03 | 0·07 | 0·37 | 0·16 |
| 6. | Denial of privileges | 0·11 | 0·20 | 0·11 | 0·31 | −0·02 | 0·16 |
| 7. | Reporting to parents | 0·68 | 0·03 | 0·14 | 0·00 | 0·11 | 0·50 |
| 8. | Lowering marks | 0·18 | 0·21 | 0·11 | −0·06 | 0·29 | 0·18 |
| 9. | Temporary eviction from class | 0·16 | 0·35 | −0·05 | 0·27 | 0·02 | 0·22 |
| 10. | Learning by heart | 0·08 | 0·15 | 0·61 | 0·16 | 0·23 | 0·48 |
| 11. | Summons to principal | 0·11 | 0·42 | 0·02 | 0·05 | 0·34 | 0·31 |
| 12. | Surprise quizzes | 0·13 | 0·02 | 0·47 | 0·05 | −0·03 | 0·24 |
| 13. | Busy work | 0·06 | 0·08 | 0·28 | 0·19 | 0·28 | 0·20 |
| 14. | Detention | 0·07 | −0·08 | 0·26 | 0·75 | 0·18 | 0·67 |
| 15. | Teacher-student conference | 0·61 | 0·10 | 0·00 | 0·06 | 0·05 | 0·38 |
| 16. | Monitor duty | 0·03 | 0·09 | 0·13 | 0·23 | 0·38 | 0·22 |
| 17. | Cleaning premises | −0·08 | −0·01 | 0·15 | 0·04 | 0·64 | 0·44 |
| 18. | Shortening recess | −0·09 | 0·13 | 0·35 | 0·25 | 0·10 | 0·24 |
| 19. | Reporting early | 0·04 | −0·08 | 0·18 | 0·64 | 0·21 | 0·50 |
| 20. | Aversive non-verbal communication | 0·00 | 0·60 | 0·09 | −0·08 | 0·05 | 0·38 |
| 21. | Verbal reprimand | 0·05 | 0·74 | 0·04 | 0·05 | 0·06 | 0·55 |
| 22. | Additional homework | 0·07 | 0·07 | 0·54 | 0·03 | 0·21 | 0·35 |
| 23. | Standing in corner | 0·09 | 0·34 | −0·01 | 0·05 | 0·33 | 0·24 |
| 24. | Changing seating | 0·05 | 0·11 | 0·26 | 0·13 | 0·18 | 0·13 |
| | % of variance | 0·08 | 0·07 | 0·06 | 0·06 | 0·06 | |
| | Label | Conferring with authority | Punitive teacher behaviour | Class related tasks | Time constraints | Extra class assignment | |

punitive teacher reaction. The factor was identified by two salient strategies, namely teacher's verbal rebuke of student and teacher's aversive nonverbal communication, with respective factor loadings of 0·64 and 0·62.

(3) The third factor, labelled, "Unpleasant Classroom Related Tasks" was marked by the following three main strategies: rote memory of school materials, additional homework assignments and surprise quizzes, loading 0·61, 0·54 and 0·47, respectively, on the factor at hand.

(4) The fourth factor, labelled "Time Constraints", was identified by two main variables, i.e., detention and arriving early at school, with item factor loadings of 0·64 and 0·62, respectively.

(5) The fifth and last factor labelled "Extra Class Assignments", include the following unpleasant extra-class assignments, i.e., cleaning school grounds and monitor duty, loading 0·64 and 0·62, respectively, on this factor.

## DISCUSSION

The major aim of the present study was to assess students' perceptions with respect to the relative severity of commonly used classroom management strategies as well as identify the specific dimensions underlying student perceptions.

The most drastic category of classroom discipline from the student's point of view involves either permanent or temporary removal from class; conferring with outside authority or threat to the student's self esteem or academic status. Following, in terms of perceived severity, are constraints and restrictions on student's time or freedom. The least drastic category involves the removal of privileges, assignment or aversive tasks to students and verbal punishment.

On the whole, student mean ratings of the severity of commonly used classroom discipline techniques and their rank order were remarkably invariant across sex and social groups, with various background variables appearing to have little effect on either the magnitude of mean ratings of specific discipline practices examined or their rank order. The data imply that males and females, on one hand, and lower- and middle-class students, on the other, tend to adopt a common "student perspective" with respect to the relative severity of certain management strategies. The same holds true for students who perceive themselves above and below average in academic attainment. The results may be attributed in part to the conformity of students to school attitudes and norms, with students tending to adopt school norms and perspectives concerning classroom discipline.

The findings of this study concur with data presented by Burns (1978) indicating that secondary school boys and girls are in almost perfect agreement with respect to their ratings of various classroom deterrents in terms of dislike. The data also agree with Burns' findings with respect to the relatively high degree of dislike of a number of classroom control techniques such as unfavourable reports home, eviction from school and ridiculing students, thus lending some cross-cultural validity to the present findings.

In the present study teachers and students were found to agree on the relative severity of the most drastic form of control (i.e., permanent eviction from school) and on the least drastic mode of control (changing seating arrangement). However, meaningful differences between teacher and student ratings were observed for about half the techniques appearing on the inventory, although the disparities were generally small. A number of discrepancies between student and teacher mean ratings are worthy of mention. Thus, whereas "personal insult to student" ranked second in terms of severity among the student group, the item ranked only in eighth place among teachers. Similarly, students perceived "parent-principal conference" as a more drastic measure than did teachers. On the other hand, teachers viewed "lowering student grade" and "aversive non-verbal communication" as a more serious form of control than did students.

The strong degree of correspondence between teachers and students in the rank order of mean classroom management strategy ratings is analogous to the high degree of correlation observed between teacher and student ratings with respect to the severity of classroom behaviour problems (Ziv and Spigel, 1969).

A factor analysis of the student ratings provided empirical evidence for a 5-factor typology of classroom control, composed of the following five control categories: (a) referral of student to school or parental authority; (b) punitive teacher reaction; (c) assignment of unpleasant classroom related task to student; (d) placing constraints on student; (e) assigning student extra-class assignments, or busy work. However, it should be noted that the proportion of total variance accounted for by each factor (ranging from 6 per cent to 8 per cent), as well as the proportion of variance covered by all five factors together (33 per cent) was quite modest, suggesting a rather tenuous factor structure for the inventory items, given the small sample size. Further work, employing larger samples, needs to be done in validating the typology across varying rating dimensions, grades and cultural contexts.

From a methodological point of view, it should be borne in mind that some of the most drastic disciplinary techniques (e.g., paddling, shaking, slapping, ear-pulling, etc.) were not included on the inventory, in view of legal prohibitions in Israel, against the use of some of the more drastic teacher response behaviour involving physical harm and abuse. Had "physical harm" been included, it probably would have been rated as one of the most severe techniques (cf. O'Hagan and Edmunds, 1982). Furthermore, the reader should bear in mind that the study was conducted on a relatively small Israeli junior high school sample of convenience so that generalisations to the population of junior high school students in general as well as to other age groups, and different cultural contexts, should be made with great caution. In addition, the management techniques were rated along other dimensions as well (e.g., perceived effectiveness, fairness, dislike, degree of anxiety evoked, etc.).

Additional research into the student's perspective on classroom discipline may provide classroom teachers with an additional source for generating appropriate decision management strategies and thus serve as a further aid to classroom managers in selecting the most appropriate measure when faced with student misbehaviour in the classroom. Indeed, by understanding the student's point of view with respect to classroom discipline and taking it into consideration when meting out sanctions, teachers may be able to handle classroom misdemeanours more effectively and consequently improve classroom discipline and control.

ACKNOWLEDGMENTS.—I thank the School of Education and Research Authorities of the University of Haifa for funding part of this research.

Correspondence and requests for reprints should be sent to Dr. Moshe Zeidner, School of Education, University of Haifa, 31999, Israel.

## REFERENCES

ANASTASI, A. (1958) *Differential Psychology* (3rd ed.). New York: Macmillan and Co.

BEILEN, H. (1959). Teacher's and clinician's attitudes towards the behaviour of children: a reappraisal, *Child Dev.*, 30, 9-25.

BURNS, R. N. (1978). The relative effectiveness of various incentives and deterrents as judged by pupils and teachers. *Educ. Stud.*, 4, 229-243.

CLARIZIO, H. F. (1976). *Toward Positive Classroom Discipline.* New York: John Wiley and Sons.

DREIKURS, R., GRUNWALD, F., and PEPPER, Ch. (1982). *Maintaining Sanity in the Classroom.* New York: Harper and Row.

O'HAGAN, F. J., and EDMUNDS, G. (1982). Pupils' attitudes towards teachers' strategies for controlling disruptive behaviour. *Br. J. educ. Psychol.*, 52, 331-340.

MCCANN, E. (1978). Children's perceptions of corporal punishment. *Educ. Stud*, 4, 167-172.

REILLY, R. R., and LEWIS, E. (1983). *Educational Psychology.* New York: Collier Macmillan.

WOLFGANG, Ch. H., and GLICKMAN, C. D. (1980). *Solving Discipline Problems.* Boston: Allyn and Bacon.

ZIV, A., and SPIGEL, Y. (1969). Different approaches towards students' behaviour problems. *Megamot behav. Sci Q.*, 16, 147-165. (In Hebrew.)

(*Manuscript received 2nd December,* 1986)

*Br. J. educ. Psychol.*, 58, 78-88, 1988

# CHILDREN'S RECIPROCATIONS OF TEACHER EVALUATIONS

BY NORMAN WORRALL
*(University of London Institute of Education),*

CHRYSOULA WORRALL AND CLAIRE MELDRUM
*(Hatfield Polytechnic)*

SUMMARY. To what extent do children and teachers form reciprocal evaluations of each other? This study followed three classes over the school year, comparing teacher ratings of children's classroom demeanour, children's interpretations of this evaluation, and children's own ratings of teachers' classroom management. The overall picture was one of girls and high achievers favouring and being favoured by their teachers. Children clearly detected the boy-girl differentiation but not differentiation by achievement group. Separate occasion analysis showed a pattern consistent with an *ab initio* causal argument for development of reciprocation on the basis of gender which must have been learned in the present classroom. For achievement group the data pattern was more consistent with a continuing than an original causal process. Among high achievers, a degree of both accurate interpretation and reciprocity developed for both boys and girls. Among low achievers, boys were generally out of touch, both in interpreting teacher evaluations and in reciprocating appropriately. Low-achieving girls showed an unexpected pattern consistent with striving to hold on to favourable, gender-based teacher evaluations from the early part of the school year.

## INTRODUCTION

FOR many years, an abiding concern in education has been how teacher biases might be influencing child outcomes. The "new look" perspective provided by Rosenthal and Jacobson's (1968) study of the self-fulfilling prophecy led to teacher behaviour being seen as the mediator rather than the cause of child outcomes with cause lying in the expectations behind that teacher behaviour. However, teacher behaviour itself was still somehow thought of as emitted as a function of type of child rather than being responsive to that child, a view represented in models as one-way causal flow from teacher to child. Beginning in the 1970s, the child came to be credited with perception of teacher behaviours, and an increasing number of studies were able to demonstrate more generally that teacher and pupil have significant influences on each other (Klein, 1971; Yarrow *et al.*, 1971; Serbin *et al.*, 1973; Rappaport and Rappaport, 1975; Feldman and Prohashka, 1979; Perry *et al.*, 1979; Feldman and Thiess, 1982).

In their original study, Rosenthal and Jacobson (1968) wrote of "feedback loops" between teacher and child, though Rosenthal's most recent formulations (Rosenthal, 1985; Harris and Rosenthal, 1985) promote essentially a one-way causal model. One-way models may indeed be justified on the grounds that they are not intended to account for each and every teacher-child influence — only the teacher expectancy effect, which could reasonably be seen as predominantly a one-way process.

Two significant accounts in the North American literature that adequately represent the idea of two-way parallel flow between pupil and teacher are those of Darley and Fazio (1980) and Mitman and Snow (1985), the latter extending the earlier formulation of West and Anderson (1976). However, the treatment is descriptive rather than leading to new analysis of the nature of the evolving relationships.

Over the past ten years or so in this country, the methodology of investigation for classroom processes has been influenced by ethnography. Among relevant examples, Benyon and Atkinson (1984) have described how children test permissible boundaries of behaviour with different teachers; Ball (1980) has analysed initial classroom encounters, and Hargreaves (1975) has given a broader account of the significance of roles and situations as understood in the classroom. While the ethnographic approach is attractive, part of the recognised trade-off for approaching some understanding of phenomena has been a sacrifice of firm, quantitative data, especially causal data.

The present study considers how effectively children can interpret their teacher's evaluation, and whether some kind of mutuality of regard may be said to exist or develop between child and teacher. These could be regarded as questions best approached from an ethnographic standpoint, although they are managed here within a normative framework, using standard research design and rating scale procedures.

While studies of teacher expectancy effects have historically concerned themselves with achievement-related dimensions, the present study has examined classroom demeanour — how "childish", "selfish", "mature", "cheeky", how much of a "troublemaker", and generally how "well behaved" particular groups of children are in class. In common with the teacher expectancy literature, we recognise the crucial fact that initial, pre-meeting expectations in both teacher and child exist without direct experience of the other at all. These expectations spring from accepted sources. For the teacher, sources may include exchanges with previous class teachers, knowing a sibling or the parents, playground gossip, systematic biases deriving from beliefs and stereotypes, and wayward but powerful ideas about "this sort of child". The child, too, probably enters the new class situation with similar ideas about "this sort of teacher". Despite their uncertain foundations, it is these initial expectations that must be the basis for the opening behaviours that teachers and pupils present in the classroom.

From the beginning, the classroom is the setting for a two-way flow of influence, child to teacher and teacher to child. It is expected that, in the earliest of such sequences, the teacher will be restricted to category-based reactions or initiations towards particular children (e.g., "average-ability sort of girl, neat and quite friendly"). However, as the acquaintance process evolves, later sequences may come to reflect appropriate individualisation and mutual understanding.

## METHOD

### *Sample*

Three classes of 9-11-year-olds were used (N=97) together with their three teachers, all female though of varied age and personality. The data for four of the children were excluded on the grounds of obvious coping difficulties. For the purpose of analysis, children were grouped by gender, and divided into upper, middle and lower ability groups on the basis of combined NFER test scores for verbal reasoning and mathematical attainment.

### *Measures*

Altogether, children and teachers provided data on 20 items covering both academic and social aspects of classroom life. These items, which were cast as 5-point, bipolar scales, were derived from descriptions of children and teachers not in the study as to the kinds of things they liked and disliked about classrooms. A principal components analysis yielded four components, including the classroom demeanour/classroom management dimensions which are the subject of this report.

Items making up classroom demeanour were: (not) selfish; (not) cheeky; (not) well behaved; (not) mature; (not) a troublemaker; (not) childish. The teacher rated each child on these six items, and in addition the child rated what he or she thought was the teacher's evaluation at that time in the year, so as to provide a "child interpretation" measure. Items making up classroom management were: fair on punishment; moody; praises good behaviour; shouts; has favourites; keeps order. Each child rated his or her teacher on these items.

### *Procedure*

The teacher withdrew from the class after introducing two of the investigators. Considerable time was spent talking children into the idea of the research and giving copious examples of how to tick rating scales. In addition, while children were completing the scales, any questions or other indications of uncertainty were dealt with. Rating forms had two possible item orders within a form, and order of forms for a given child was randomised. Teachers themselves completed the rating for each child in the staff room or that evening at home. For the purposes of the present report, there were four measuring occasions: Day 1; Day 93 (near the end of the first term); Day 214 (middle of the spring term), and Day 308 (towards the end of the summer term).

## RESULTS

In the analysis of results, school class typically emerged as a main effect, so that amount of difference between one measure and another varied from class to class. However, as noted in Table 1, only once did class interact with gender or achievement, so that for the most part it is possible to present data in terms of 93 children rather than three classes.

### *The causal argument from mean differences*

If one group of children, say girls, both appraise the teacher's classroom management more highly and are themselves appraised more highly by teacher immediately from Day 1, then it may be inferred that such mutual liking must be based on predominantly pre-experiential factors. That is to say, although "primitives" such as dress, voice and physical appearance will have begun contributing to impressions, it is the pre-experiential factors — stereotypes and beliefs about boys and girls, recall of "how it was last year" — which will be dominant. Any similarities in mutual regard that may be found in the data at Day 1 will, therefore, be interpretable as simple correspondence rather than causal reciprocation arising from direct experience in the classroom.

On the other hand, the possibility of a causal argument is at least opened by an *asynchronous* pattern in which, say, teacher preferences for some group of children exist at Day 1 without any return favouring being expressed by that group for the teacher, and then at a later date both the teacher and those same children emerge as showing mutual favouring. In this particular example, an additional inference would be that the predominant direction of causation was from teacher to child.

(a) *Gender-based reciprocation.* It can be seen from tracing the pattern of F values in Table 1 that the foregoing account fits the present data quite well. At Day 1, teachers immediately rate girls better than boys on classroom demeanour ($F=5{\cdot}75$). Moreover, the Day 1 data which show whether teacher-girl reciprocation already exists, namely, children's regard for the teacher, shows no such gender differentiation ($F=0{\cdot}26$). Thus the Day 1 pattern indicates (a) that teacher-pupil correspondence is not immediately present on a gender basis, so that the asymmetric conditions for eventual demonstration of reciprocation are present, and (b) that the preponderant direction of any causal flow must be from teacher to child.

TABLE 1

SUMMARY OF F VALUES FOR THE THREE CLASSROOM MEASURES OVER THE SCHOOL YEAR

| Measure | Design Factor | Day 1 | Day 93 | Day 214 | Day 308 | Linear Trend |
|---|---|---|---|---|---|---|
| Teacher Rating | Gender | 5·75** | 2·73 (*) | 7·28*** | 3·89* | 0·03 |
| of Child | Achievement | 4·21** | 2·78 (*) | 9·60*** | 5·59** | 0·56 |
| Child's | Gender | 1·53 | 3·27 (*) | 5·41* | 1·06 | 1·66 |
| Interpretation | Achievement | 0·84 | 0·71 | 0·46 | 0·46 | 1·19 |
| Child's Rating | Gender | 0·26 | 4·72* | 15·16*** | 10·76*** | 8·21** |
| of Teacher | Achievement | 4·17** | 1·55 | 0·35 | 1·77 | 1·29 |

*Notes:* *** P < 0·001; ** P < 0·01; * P < 0·05; (*) P > 0·05 < 0·10. For no occasion was there a significant interaction between Gender and Achievement Grouping, or any involving the Class factor. In the linear trend analysis, there was a Gender × Achievement interaction for Teacher Ratings, and a Gender × Achievement × Class interaction for Child Ratings of Teacher.

It then becomes fairly clear from Table 1 that pick-up of teacher differentiation — which is itself consistent over the year — appears in the children's interpretations around Day 93 (F = 3·27), while reciprocation has also appeared reliably by this date (F = 4·72). Since the direction of effect is for girls to rate their teacher higher than boys, this is a precise reciprocation, because (a) the teachers themselves initially favoured girls and continued to do so; (b) once initiated, girls' favouring of their teachers also remains throughout the year. Consistent with this pattern, there is in Table 1 a reliable linear trend in the children's evaluations of the teacher (F = 8·21). This trend is to boy-girl separation over the year, with boys' view of the teacher deteriorating.

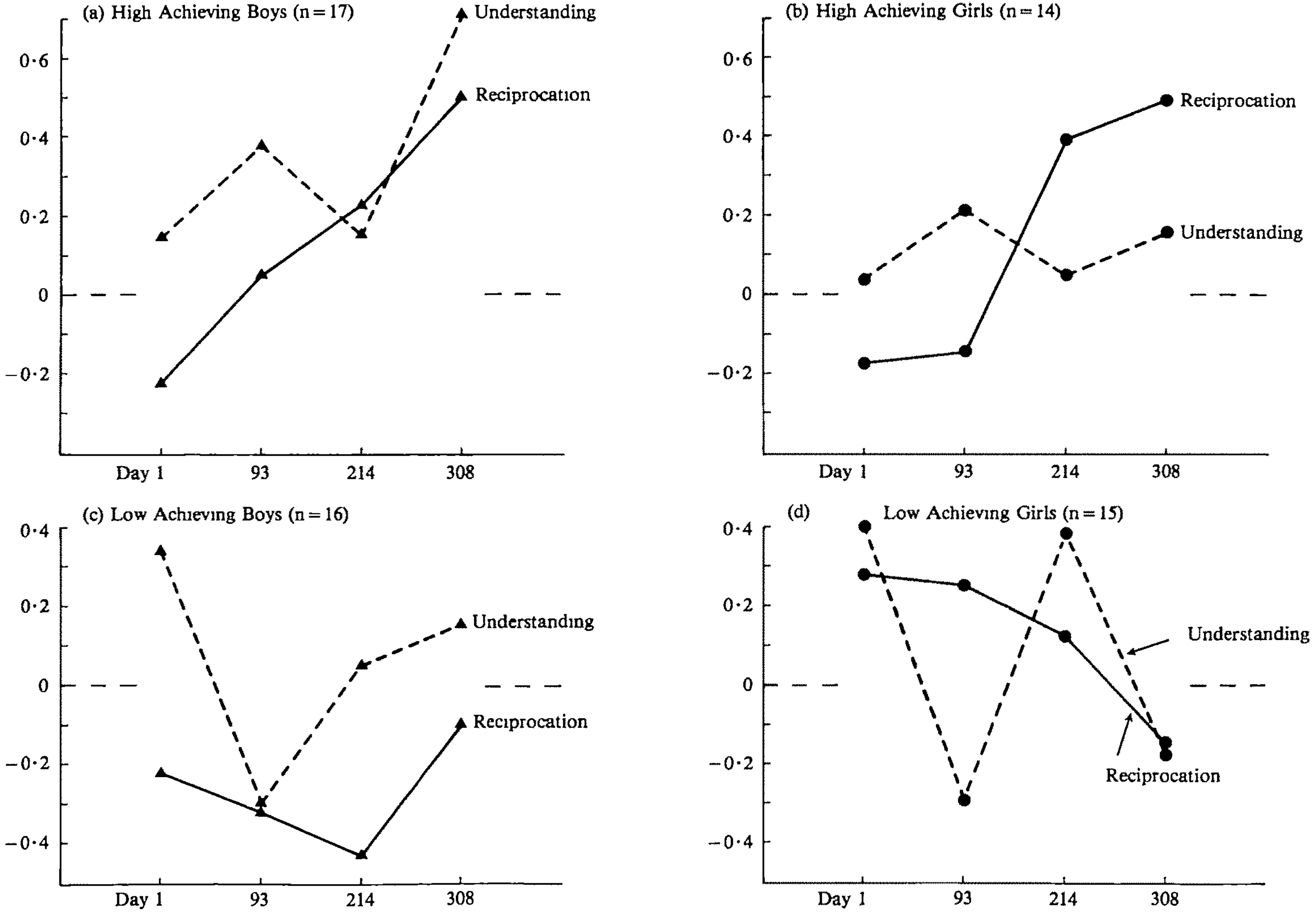
(b) *Achievement-based reciprocation.* Parallel patterns may be sought on the basis of children's achievement groupings. Referring again to Table 1, teachers' more favourable view of higher achieving children's classroom demeanour is evidence from Day 1 (F = 4·21) and is maintained throughout the year. The difference here, however, as Table 1 shows, is that children also rate their teachers differentially on the basis of achievement group immediately from Day 1 (F = 4·17). Because higher achieving children already favour their teacher on classroom management at Day 1 an *ab initio* causal argument based on specific classroom experience seems ruled out. These teachers are already disposed to favour brighter children and the brighter children disposed to favour their teacher. As soon as each approximately identifies the other the disposition is implemented, with the classroom merely being the setting within which the mutual favouring is released. It is worth noting that, while children's return evaluations of their teacher do show differentiation by achievement group at Day 1, its consistency is not generally impressive across measuring occasions (Row 6 in Table 1). However, separation between high and low groups is quite well maintained.

### *The causal argument from covariation*

A second causal argument can be developed, not now from asynchronous differences between means, but from actual covariation over the year between teacher and child mutual ratings. Correlation rather than regression is the measurement model of choice, since both reciprocation and simple correspondence imply two variables of equal causal status, rather than one dependent and one independent variable.

FIGURE 1

CORRELATION TRENDS OVER THE SCHOOL YEAR FOR THE FOUR GENDER-BY-ABILITY SUBGROUPS



*Note:* The Understanding graphs show correlations between teacher regard on classroom behaviour and children's interpretations of that regard. The Reciprocation graphs show correlations between teacher regard for children on classroom behaviour and children's regard for teachers

Figure 1 presents correlational data for the four groups of maximum contrast: high achieving girls (N = 14); high achieving boys (N = 17); low achieving girls (N = 15), and low achieving boys (N = 16). In each panel of the figure, two correlation trends are shown. The first is the correlation between teacher rating of child and child rating of teacher, and is labelled "reciprocation". The second is the correlation between teacher rating of the child with the child's interpretation, so as to give the graph labelled "understanding".

For the high achieving boys it can be seen that there is a clear trend to improvement in the reciprocation index over the year, from −0·22 at Day 1 to 0·50 by the last day. In the same panel, the understanding trend is less striking but is always clearly positive and upwards. Altogether, these high achieving boys present the best picture of teacher-pupil empathy for any of the four groups. The corresponding panel for high achieving girls shows that the reciprocation correlations are similar to those of the boy counterparts, improving from −0·17 to 0·49. However, the understanding correlations are less impressive, remaining at 0·20 or less all year.

The remaining panels of Figure 1 present a considerably less favourable picture for low achieving girls and boys. In particular, for low achieving boys the reciprocation correlations are always negative, so that throughout the year the dispositions of individual boys to their teachers were generally the reverse of what teacher was believing about them. Their correlations for understanding offer little insight into possible reasons, since on three out of the four measuring occasions they are actually positive. Finally, and perhaps most remarkably, there are the reciprocation correlations for the low achieving girls. Whereas the low achieving boys had shown a consistent absence of mutuality with their teachers, these girls present a systematic picture of becoming disposed to their teacher in a way which is progressively opposite to their teachers' return dispositions. The correlations for understanding again offer no insights here, being too erratic for useful interpretation.

*Further diagnostics.* The unexpected correlation patterns in the low achieving data suggest a return to the constituent measures to understand why. As Figure 2 shows, the general pattern is for teachers' ratings of children to be the most favourable, followed by children's interpretations and finally the children's return rating of the teacher. This is seen for high achieving boys and girls and even for low achieving boys. The interesting exception to this pattern is the case of low achieving girls.

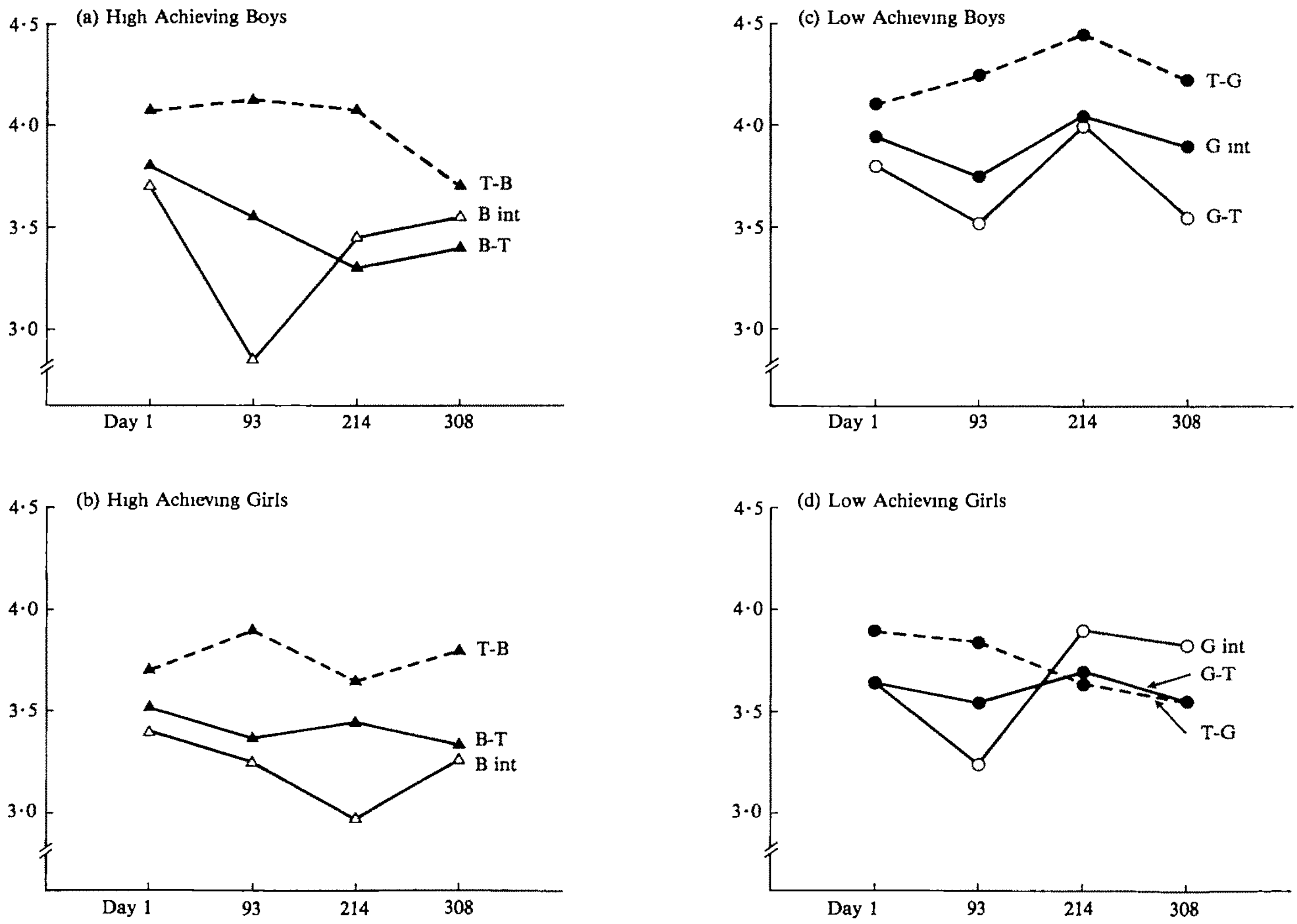
Teachers themselves evaluated high achieving girls and boys nearly identically. Furthermore, the low achieving boys show about the same separation among the three measures as do the two high achieving groups. It is, thus, the low achieving girls who are indeed the odd group out in Figure 2, since their child-rating-of-teacher graph closes on and at one point crosses the teacher rating graph. A similar crossover occurs even more markedly for their interpretation ratings. In other words, relative to the other three groups, the low achieving girls show a degree of both over-interpretation and over-reciprocation.

## DISCUSSION

### *Simple correspondence or causal reciprocation?*

The first question concerns whether some correspondence could be said to exist between pupils or some subgroup of pupils and their teachers on this classroom demeanour/classroom management dimension. If yes, a second question concerns whether the causal basis for this correspondence may be found in the present classrooms or has to be assigned to outside factors. If a causal basis within the

FIGURE 2
MEAN TRENDS OVER THE YEAR FOR THE THREE MEASURES



*Note:* Graphs show teachers' regard for boys (T-B) or girls (T-G) on classroom behaviour; boys' or girls' interpretations of teachers' regard (B int; G int), and boys' or girls' regard for their teachers on classroom management (B-T; G-T).

classroom context can be established — and thereby true reciprocation — then a third question concerns what may be said about the nature of this reciprocation.

Evidence for simple correspondence can most readily be seen in the correlation data for the two high achieving groups, where by terms 2 and 3 clear teacher-pupil mutuality of regard was appearing. The claim that this correspondence was causal depends on the asynchronous appearance of differentiation in the teacher and child data, which may be summarised as follows. The Day 1 picture was one of teacher favouring of girls but no sensing of this, nor any return favouring from girls themselves. By Day 93 the continuing teacher favouring was now being picked up by the children, and also being returned in girls' better ratings of their teachers. Moreover, in addition to these fixed-occasion data, there was a trend to increased differentiation in boys' and girls' return evaluation of their teachers over the year. Thus Table 1 provides a four-way structure for a gender-based causal argument.

However, the correlation data in Figure 1 modify this basic finding in an important way by showing that once we look *within* subgroups, rather than merely comparing their mean levels, then reciprocation in the explicit sense of one-to-one teacher-child matching developed only for high achieving girls and boys. Thus it was the high achieving groups that recognised the differentiation being made by the teacher *and* reciprocated appropriately, while the low achieving groups' interpretations merely recognised the differentiation and did not go on to reciprocate appropriately.

From an achievement group standpoint, the possibility of developing an *ab initio* causal argument from the mean level data in Table 1 seemed weaker. Achievement-based correspondence was something that was either brought into, or very rapidly activated in the classroom, since it was sufficiently evident in the Day 1 ratings of teachers and children to yield differences in group means. Nevertheless, the simultaneous fact that high but not low achieving children moved from negative to positive *correlations* over the year showed that individualisation of this early, "coarse" differentiation was also occurring.

We would therefore characterise the achievement-based argument as being not so much about original as continuing causal processes. We also have in mind here the view of Cooper (1979) that the more common influence of teacher on child is not so much one of changing as of maintaining, where teachers' early impressions are largely accurate but go on to maintain or constrain what the child may then achieve. Within present terms, this is equivalent to saying that once achievement level is coarsely identified then reciprocation merely unfolds in a group-appropriate manner, and that the correlation trends are diagnostic of this unfolding process.

A last point on achievement groupings is that because teachers' achievement-based differentiations both began and remained strong throughout the year, while differentiation from children became suspect after Day 1, the maintaining causal flow in the *correlational* reciprocation that developed may be taken as again predominantly from teacher to child.

To summarise, while gender-based reciprocation was learned or quite possibly re-learned in the three classrooms, achievement-based reciprocation merely developed from a pre-existing base. In both cases, and perhaps unremarkably, the preponderant causal flow was from teacher to child.

### *The position of low-achieving children*

(a) Were children detecting achievement-related messages from the teacher? It is fairly clear from the correlational data that high achieving children at least developed a reasonably accurate understanding of teacher regard of classroom demeanour. Thus, regard of some kind was being expressed by the teachers, although it does not follow that opposite messages were being conveyed to high achieving and low achieving children. And how, in any case, was achievement group being translated by teachers into an evaluation of classroom demeanour? One possibility is that low achieving children suffer from a general negative halo. Heller and White (1975) and Sherry *et al.* (1980) have noted how teachers make more disapproving comments to lower achievers, most of these comments being on the dimension of interest here — classroom demeanour.

A second reason could be that low achieving children are in fact naughtier, in so far as lower interest levels and attention spans, as well as the need for an alternative class image, make them more open to distracting and being distracted. However, the evidence for this is much clearer for low achieving boys than for low achieving girls (e.g., Good and Brophy, 1974, Table 5.2).

Concerning children's own perceptions, 10-year-olds are quite capable of responding to gender-tagged outflow. As the data showed, they are also quite capable with women teachers at least of reciprocating vigorously on a gender basis. It appears, however, that up to the age of 11 or 12, children are to some degree confused about what constitutes "ability" and what "effort", believing, for example, that those who try harder are somehow more able (Nicholls, 1978; Stipek, 1981). Even if they know what ability is, what 10-year-olds still do not know very clearly is whether they *themselves* are "high" or "low" ability. This especially applies when teachers do not signal such differentiation (Brattesani *et al.*, 1984). In addition, even when teachers do put out differential messages they may translate low ability into poor effort for the sake of palatability, and thereby compound the children's non-understanding.

Since the children have no proper concept of themselves in terms of *ability*, they cannot respond to specific, ability-tagged messages from the teacher. Thus it becomes less surprising that children's interpretations in Table 1 showed no achievement-based differences. But this account does not allow for the correlation data being so responsive to achievement grouping. The point here, we suggest, is that regardless of whether children can conceptualise ability or receive and respond to ability-tagged messages from the teacher, the fact remains that some of these children *are* brighter than others. Because they are brighter, when we as researchers partition them into achievement groups and examine how accurately as individuals within groups they can interpret or reciprocate, then cognitive superiority will make for better child-teacher matching amongst the highs. Hence, the superior correlations for the present high achieving children in Figure 1.

(b) Can children's ratings, particularly of low achievers, be taken at face value? If any group in the present study had cause to distort, it would presumably be the "disadvantaged" low achievers. The mean level data for low achieving boys were ordered in a very similar way to the data for the high achieving groups, so that there was no obvious distortion there. However, as we saw, some systematic bending of ratings was occurring with low achieving girls, since relative to the other groups they interpreted how teachers acted towards them excessively favourably, and went on to relatively overestimate their teachers. We see this as evidence for low achieving girls striving to hold on to something they sense as slipping away after the first term of the new class. They overvalue the teacher herself, and deny in their interpretations the relatively low regard she has for their classroom demeanour. It is a pattern that makes them quite different not only from low achieving boys, but also from other girls who happen to be high achievers. It could hardly be maintained, therefore, that low achieving girls are "out of touch" with messages coming from their teacher. On the contrary, their data pattern is consistent with a picture of effortful striving in response to such messages.

*Concluding Comment.* It is evident that even on the single dimension of classroom demeanour, no study involving only three classes of children can hope to claim what is the case across primary classrooms generally. It is especially not wished to claim here that gender-based differences will always be found and that these are somehow the "fault" of the teacher, whereas achievement-related differentiations will merely be an inevitable elaboration of the status quo.

As noted earlier, school class typically emerged as a reliable main effect. However, class typically did not interact with gender or achievement, so that the generality of theoretically interesting findings was not compromised. The fact that findings essentially replicate across the three classes is also some defence against spuriousness which threatens all attempts at causal argument, particularly within non-experimental studies such as this one.

Much of the strength of the present study lies in the fact that it was longitudinal and thus could examine the stability of constructs over different occasions. In the existing literature, single-occasion classroom studies are by far the most common and correspondingly vulnerable to the vagaries of the particular occasion. It is clear from Table 1 or Figure 1, for example, that a single-occasion study would have yielded a quite different data complex and subsequent argument had it been run only at Day 1 or 93 or 214 or 308.

ACKNOWLEDGMENTS.—We are grateful to the SSRC (now ESRC) for supporting this research, to Charles Owen for assistance with data analysis, and to the teachers and children for willingly taking part.

Correspondence and requests for reprints should be addressed to Dr. Norman Worrall. University of London Institute of Education, Department of Educational Psychology and Special Educational Needs, 24-27 Woburn Square, London, WC1H 0AA.

## REFERENCES

BALL, S. (1980). Initial encounters in the classroom. In WOODS, P. (Ed.), *Pupil Strategies*. London: Croom Helm.

BENYON, J., and ATKINSON, P. (1984). Pupils as data gatherers: mucking and sussing. In DELAMONT, S. (Ed.), *Readings on Interaction in the Classroom*. London: Methuen.

BRATTESANI, K., WEINSTEIN, R., and MARSHALL, H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *J. educ. Psychol.*, 76, 236-247.

COOPER, H. (1979). Pygmalion grows up: a model for teacher expectation communication and performance influence. *Rev. educ. Res.*, 49, 389-410.

DARLEY, J., and FAZIO, R. (1980). Expectancy confirmation processes arising in the social interaction sequence. *Am. Psychol.*, 35, 867-881.

FELDMAN, R., and PROHASHKA, T. (1979). The student as Pygmalion: effect of student expectation on the teacher. *J educ. Psychol.*, 71, 485-493.

FELDMAN, R., and THIESS, A. (1982). The teacher and student as Pygmalions· joint effects of teacher and student expectations. *J. educ Psychol.*, 74, 217-223.

GOOD, T., and BROPHY, J. (1974). *Teacher-Student Relationships: Causes and Consequences*. New York: Holt, Rinehart and Winston.

HARGREAVES, D. (1975). *Interpersonal Relations and Education*. London: Routledge and Kegan Paul.

HARRIS, M., and ROSENTHAL, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychol. Bull*, 97, 363-386.

HELLER, M., and WHITE, M. (1975). Rates of teacher approval and disapproval to higher and lower ability classes. *J. educ. Psychol.*, 66, 796-800.

JACKSON, P., SILBERMAN, M., and WOLFSON, B. (1969). Signs of personal involvement in teachers' personal descriptions of their students. *J. educ. Psychol.*, 60, 22-27.

KLEIN, S. (1971). Student influence on teacher behavior. *Am. educ. Res. J.*, 8, 403-421.

MITMAN, A., and SNOW, R. (1985). Logical and methodological problems in teacher expectancy research. In DUSEK, J. (Ed.), *Teacher Expectancies*. Hillsdale, NJ: Erlbaum.

NICHOLLS, J. (1978). The development of the concepts of effort and ability, perception of academic attainment and the understanding that difficult tasks require more ability. *Child Dev.*, 49, 800-814.

PERRY, R., ABRAMI, P., LEVENTHAL, L., and CHECK, J. (1979). Instructor reputation: an expectancy relationship involving student ratings and achievement. *J. educ. Psychol.*, 71, 776-787

RAPPAPORT, M., and RAPPAPORT, H. (1975). The other half of the expectancy equation: Pygmalion. *J. educ. Psychol.*, 67, 531-536.

ROSENTHAL, R. (1985). From unconscious experimenter bias to teacher expectancy effects. In DUSEK, J. (Ed.), *Teacher Expectancies*. Hillsdale, NJ: Erlbaum.

ROSENTHAL, R., and JACOBSON, L. (1968). *Pygmalion in the Classroom*. New York: Holt, Rinehart and Winston.

SERBIN, L., O'LEARY, K., KENT, R., and TONICK, I. (1973). A comparison of teacher response to the preacademic and problem behavior of boys and girls. *Child Dev.* 44, 796-804.

SHERRY, L., ARMSTRONG, S., and ALGOZZINE, R. (1980). Teachers' perceptions and classroom interactions. *Psychol. Rep.*, 46, 535-540.

STIPEK, D. (1981). Children's perceptions of their own and their classmates' ability. *J. educ. Psychol*, 73, 404-410.

WEST, C., and ANDERSON, T. (1976). The question of preponderant causation in teacher expectancy research. *Rev. educ. Res.*, 46, 613-630.

YARROW, M., WAXLER, C., and SCOTT, P. (1971). Child effects on adult behavior. *Dev. Psychol.*, 5, 300-311.

(*Manuscript received* 16*th September*, 1986)

*Br. J. educ. Psychol.*, 58, 89-102, 1988

# SECTIONAL DRAWINGS FROM SCIENCE TEXTBOOKS: AN EXPERIMENTAL INVESTIGATION INTO PUPILS' UNDERSTANDING

BY HILARY CONSTABLE
*(School of Education, University of Leeds),*
BOB CAMPBELL
*(Department of Education, University of York)*
AND RON BROWN
*(Faculty of Education, New College, Durham)*

SUMMARY. This paper examines aspects of children's understanding of illustrations used as an adjunct to learning. It reports an experiment designed to determine whether first-year secondary school pupils could identify the cut surfaces of objects in six biological illustrations taken from commonly used textbooks.

The results show that few children were able to perform the tasks correctly but that the illustrations were not equally difficult. Picture analysis indicated that not only the features of the object depicted but also the number and type of pictorial conventions employed posed significant difficulties.

It is concluded that the findings have implications both for theories of perceptual development and for the role of illustrations in teaching. It is suggested that teachers should regard the use of illustrations as learning aids as problematical and consider the need to teach *about* illustrations.

## INTRODUCTION

TEACHERS and publishers of educational materials need to select illustrations, but it is still not clear on what basis they should do so. The question of criteria for selection has been highlighted and in some respects the field has been much researched. Some recommendations for the selection of pictures have been made (Dwyer, 1972; Blaine, 1982) and children's preferences have been charted (Campbell, 1976; Holliday *et al.*, 1977; Petterson, 1982; Brown, 1985) but, in spite of this, a coherent synthesis of the role of pictures in learning has been slow to emerge. The task of delineating criteria for the selection of illustrations exists in an ambiguous environment. While there is general agreement that illustrations are valued in teaching, some reservations as to this value have been expressed and several pieces of experimental research lend weight to such views (Samuels, 1970; Lippman and Shanahan, 1973; Concannon, 1975).

The use of pictures as components of teaching and learning packages can be justified on two grounds. Firstly that illustrations assist learning and secondly that learning to use illustrations is a valuable educational aim in itself. Furthermore, in certain disciplines, for example geography and science, it can be argued that learning to use illustrations is part of essential subject knowledge.

Fleming (1979), in reviewing the work of the previous 20 years, notes that research has improved in that it has become more analytical and that advances in psychology have contributed to our understanding. In spite of this, much of the work seems to have only limited relevance to our concern: the value of illustrations in assisting learning in science. One much researched aspect of the use of pictures in education is whether and to what degree children can recognise the objects depicted in illustrations. Numerous investigations have tested this (Elkind *et al.*, 1964; Friedman and Stephenson, 1975). The substance of these tests has been identification, the success of which has been judged by the ability of pupils either to name or to match objects. Such an approach does not take into account the fact that pictures contain information about the relationships between the parts depicted. To determine whether a subject can "read" such information necessitates an approach other than one of simply naming the object being depicted. Serpell (1976) appreciated this point and in his studies on depth perception he relied not on a subject's ability to name parts but on his understanding of the relationship of the parts in the whole picture.

What sense a child makes of an illustration is an important question. An analysis which bears on this is provided by Duchastel (1978). He discriminates the functions of illustrations into three broad categories: attentional, retentional and explicative. It is generally thought that it is in this latter explicative role that illustrations are particularly powerful learning aids.

> "They explain a given aspect of the topic being presented or they add something which is not clearly expressible in words" (p. 38).

In these circumstances it is vital that the pupil makes the same sense of an illustration as the teacher and that the illustration contains the information which allows the child to do this. When a biology teacher selects an illustration of, for example, the pond organism hydra, he or she knows what is important about hydra and thereby has a cognitive structure which necessarily shapes his or her perception of the illustration. Whether the pupil makes the same sense when presented with the illustration is unknown. This vital piece of information is necessary both for the selection of appropriate illustrations and for the tailoring of their use. This question of interpretation is fundamental and its neglect is a stumbling block to the synthesis of our understanding of the role of pictures in learning.

Explanations offered to account for the progress which children make in picture interpretation can be conveniently divided into those which relate to general development, such as cognition (Beagles-Roos and Greenfield, 1979), language (Hale, 1971), or psycho-motor skills (Nesbit, 1981), and those which postulate abilities that are specific to the process of picture interpretation. In this latter category is the somewhat contentious theory that some advances can be attributed to learning of pictorial conventions (Gombrich, 1965; Nicholson and Seddon, 1977; Luczoz, 1982). The contentiousness has its roots in disagreement about the nature of pictures themselves. Gibson (1954) and Wittgenstein (1961) argue that pictorial representation depends upon visual correspondence, or "likeness" with the object depicted, whereas Godman (1968) and Arnheim (1969) hold the view that it depends upon arbitrary symbols or conventions. The former position sees development in picture interpretation skills as the increasing ability to make intelligent judgments about pictures which have imperfect fidelity with the object depicted. The latter explains progress by postulating a growing vocabulary of pictorial symbols and conventions. Implications of this work are that some pictorial conventions need more learning than others and some illustrations pose difficulties because they make greater use of conventions than others. The essence of such conventionalisation is economy, in that a great deal might be represented by a few marks, although other features such as topographical distortion do play a part. Kennedy (1974) developed a system of picture analysis which takes the elements of any representation as contour, cast shadow, highlighting and variation of inclination of surface. He goes on to show that line alone can carry all this information in that it can represent occluding bounds and edges, both of these with and without background surfaces, and concave and convex dihedral angles. From this information the viewer can "fill in" the missing information such as pigmentation and illumination. Since the viewer makes such a large contribution to such representations the need for familiarity is considerable, particularly so when, as in some drawings, a single line may be performing a number of functions, for example both change of direction and change of pigmentation.

School science texts make a great deal of use of highly conventionalised illustrations (Holliday, 1979; Constable, 1983). One type of illustration which occurs frequently in many science books is the sectional drawing. Such drawings may be conventionalised in a number of ways and it could be argued that until the viewer has internalised the conventions there is considerable scope for misperceptions to be made. These could arise from several causes, such as confusion between lines intended to bound the cut surfaces and lines indicating layers within the material. Further confusion might be caused by the absence of shading or cast shadow to describe the surface. It was the purpose of the study reported here to illuminate such potential difficulty by investigating the extent to which pupils could interpret such conventionalised illustrations. In contrast to previous studies of picture perception the study is characterised by the following features:

(1) It is concerned with the abilities of pupils of secondary school age.

(2) It deals with illustrations which can be considered as highly conventionalised.

(3) It employs illustrations taken from published texts aimed at the age group in question.

(4) It focuses on one aspect of an illustration: cut surface.

## EXPERIMENTAL DESIGN

In designing the study it was considered important that any task which was set did not merely test recall of knowledge since this would confuse the information conveyed by a pictorial representation with existing cognitive structures. The essential element in the design was thus a task which aimed to discover whether children could identify the cut surfaces in sectional drawings of objects with which they were unfamiliar.

In order for the results of such a task to be meaningful it was considered necessary to investigate two factors:

(1) The extent to which children could understand the terms to be used when applied to a familiar object.

(2) The ability of children to understand an illustration of an object with which they were familiar.

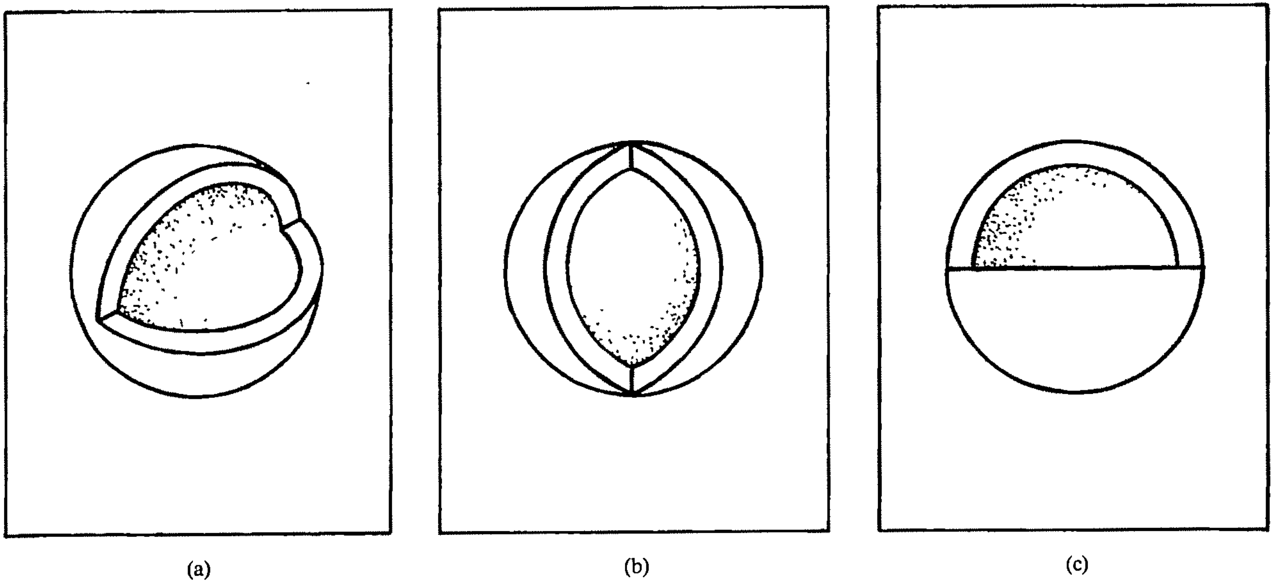
A ball was chosen as an object with which all children were likely to be familiar and two tasks were designed. The first of these was to determine if children could identify correctly the outside, inside and cut surface of a cut ball while the second was designed to establish if they could identify the cut surface in illustrations of the same ball (Figure 1).

It was considered that if both these tasks could be completed satisfactorily, then two factors would have been eliminated. These were the possibility that pupils did not understand the concept of cut surface or share the experimenter's vocabulary and the possibility that pupils could not identify the cut surface in an illustration of a familiar object.

FIGURE 1

BALL ILLUSTRATIONS

The ball illustrations presented to pupils. These were presented individually on A4 sheets of paper. They are shown here, for convenience, in reduced form but the figure shows the relative proportions of illustration and sheet.



(a) (b) (c)

Thus a sequence of tasks emerged: ball task, ball illustration task, book illustration task. In arriving at such a sequence it was recognised that, while the ability to recognise features of an illustration of an object would be isolated from a need to be able to recognise the object itself, there was a possibility that the recognition of the cut surfaces in the book illustration could be influenced by presenting pupils with the ball illustrations. To cater for the possibility of this influence, the procedure to be adopted for the investigation required that only half the number of pupils in the sample population be presented with the ball illustrations. It was considered that this procedure would thus be able to establish if pupils understood the terminology, whether they could identify the required features of the book illustrations and if the presentation of the ball illustrations had any influence on this latter result.

Illustrations for use in the experimental study were selected in several stages. Firstly, a range of commonly used lower school science texts and curriculum materials was examined (what these texts were is reported separately; Constable, 1983) and a collection of all illustrations showing cut surface was made. Secondly, criteria for selection were generated and applied. These were that a request to identify cut surfaces should be one that pupils could reasonably be expected to fulfil, that the experimenters should be clear when pupils had done so, and that it could be reasonably assumed that the illustration would NOT be familiar to the pupils. Six illustrations were chosen by this process. These are presented in Figures 2 and 3 and are hereafter referred to in terms of the biological materials they represent: Alveoli, Spirogyra, Hydra, Blastula, Fish and Uterus. (Note that these are presented in a reduced size from those used in the study.) The sources of these illustrations are given in Appendix 1.

## METHOD

The tasks were given to 12 groups of 12 pupils chosen from the mixed ability, first year classes of three 11-18 comprehensive schools. All the pupils in each group completed a series of either two or three tasks. All pupils were given a preliminary task with a cut squash ball; half had the intermediate task with illustrations of the squash ball and all pupils completed a task with three illustrations from textbooks.

A total of 144 folders of illustrations were made up, one for each pupil. The contents were allocated according to a matrix in which boys or girls, and inclusion or not of ball illustrations were present in all combinations. In addition, the order of illustrations was presented in all permutations.

For the preliminary task each child was issued with a box containing a squash ball with a quadrant cut out and asked to indicate the outside, inside and cut surfaces by means of coloured stickers. For the illustrations task each child was given a folder of illustrations and asked to colour in the cut surfaces of each. As a cross-check on the experimenters' view that the biological illustrations and the objects they represented were unknown to the pupils, they were asked to try to identify the objects.

The possible colouring patterns for each area on each illustration were that a pupil could colour an area representing a cut surface, could colour an area not representing a cut surface, or could fail to colour an area representing a cut surface. It was thus possible to devise a simple scoring scheme: each area in each illustration was scored as either under-coloured (U), correctly coloured (C) or over-coloured (O). A global score for the correctness of one child's work with one illustration was made up of scores of U, C or O for each area of the illustration. No difficulty was experienced in applying this scheme.

FIGURE 2

Group 1 Book Illustrations

Group 1 book illustrations presented to pupils comprising of (d) Alveoli, (e) Spirogyra and (f) Hydra. These were presented individually on A4 sheets of paper. They are shown here, for convenience, in reduced form. As each illustration presented was an exact duplicate of the original book illustration, the illustrations are not all to the same scale.

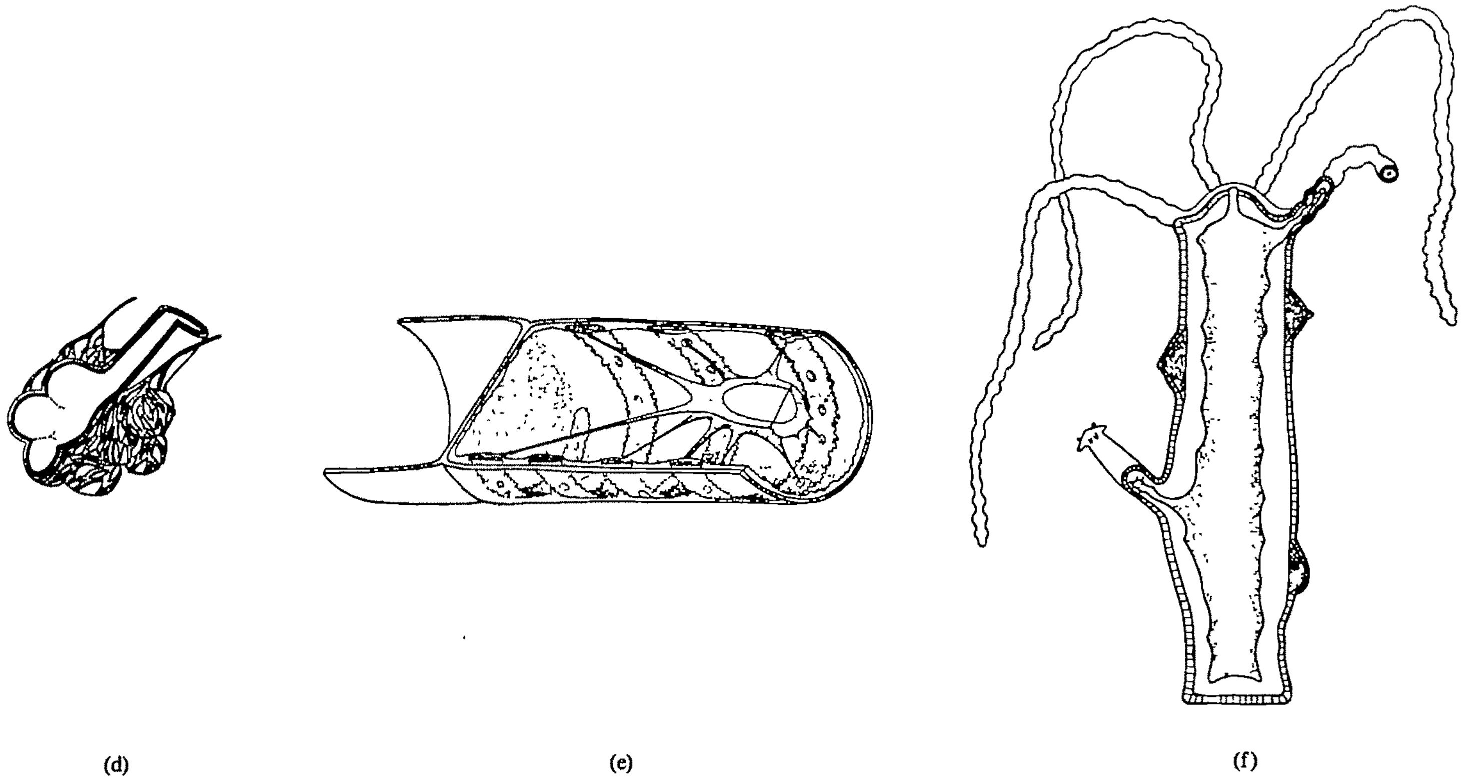
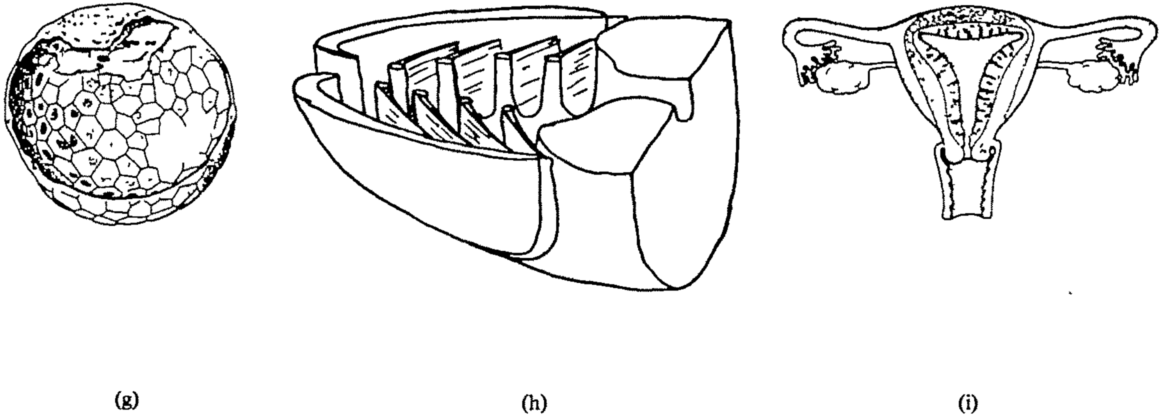


FIGURE 3

GROUP 2 BOOK ILLUSTRATIONS

Group 2 book illustrations presented to pupils comprising (g) Blastula, (h) Fish and (i) Uterus. These were presented individually on A4 sheets of paper. They are shown here, for convenience, in reduced form. As each illustration presented was an exact duplicate of the original book illustration, the illustrations are not all to the same scale.



## RESULTS

*Balls*

All 144 children in the study identified the surfaces of the squash ball correctly. This established that there was a common vocabulary between experimenters and subjects and that the pupils had some understanding of the concepts of outside surface, inside surface and cut surface.

*Ball illustrations*

Of the 216 attempts by the 72 children given the task of colouring the cut surfaces of the ball illustrations 211 were judged to be correct. No pupil completely failed to perform the task. Thus the pupils demonstrated an ability to identify the cut surfaces on these illustrations.

*Book illustrations*

(a) In marked contrast to the ball illustrations only 73 out of 432 attempts to identify cut surfaces were judged to be correct.

(b) Seeing ball illustrations before book illustrations neither enhanced nor depressed pupils' ability to colour book illustrations correctly. Out of a total of 216 attempts at book illustrations tackled after ball illustrations only 37 were correct compared with 36 of 216 correct when no ball illustrations were presented.

(c) Order of presentation of the illustrations, sex, school and teaching group each made no difference to the pupils' ability to colour the illustrations correctly.

(d) From the 432 pupil/illustration presentations only 91 attempts were made to name the structures in the illustrations. Decisions on the correctness or otherwise of answers were necessarily somewhat arbitrary and if anything generous as it was felt that pupils might not have been able to cite names even though they may have recognised a structure. Even so, only 14 of the attempts were judged to be correct. There is thus ample evidence that the pupils were not familiar with the structures depicted.

(e) Pupils did not score equally on the different book illustrations (Table 1). One illustration (Alveoli) was much more frequently coded correctly than any other. The Blastula and Fish illustrations were coded correctly only a few times. Spirogyra, Hydra and Uterus were each coded correctly by only one pupil.

(f) The size of the illustration did not influence success in colouring.

On the assumption that a low global score indicated that an illustration was difficult to colour code and a high score that it was easy, the detail of the pupils' colouring was re-examined with a view to gaining some insight to the possible causes of the difficulties. This inspection revealed that for each illustration some patterns of colouring were more common than others. By analysing the features of the illustrations and attempting to match these with colouring patterns, two aspects of the illustrations were postulated as possible causes of the coding difficulties. One aspect related to the features of the object portrayed: the number of layers represented; a change in direction of the plane of cut; non-saggital sectioning, width of cut surface and spatial discontinuity in the cut surface. A second was the means by which the object was depicted, in particular, the use of line and shading.

Looking first at the multiple nature of the cut surface, the evidence suggests that the illustrations of Spirogyra, Hydra and Uterus posed some difficulty. For

TABLE 1

PUPIL CODING OF DIFFERENT BOOK ILLUSTRATIONS

| Illustrations | | | Number of Attempts | Number of Successful Attempts | % Success | Order of Success |
|---|---|---|---|---|---|---|
| Group | Code | Name | | | | |
| 1 | d | Alveoli | 72 | 54 | 75 | 1 |
| 2 | h | Fish | 72 | 9 | 12·5 | 2 |
| 2 | g | Blastula | 72 | 7 | 9·7 | 3 |
| 1 | e | Spirogyra | 72 | 1 | 1·4 | 4 |
| 1 | f | Hydra | 72 | 1 | 1·4 | 4 |
| 2 | i | Uterus | 72 | 1 | 1·4 | 4 |
| | | TOTALS | 432 | 73 | 16·9 | |

Spirogyra, all but two of the 72 pupils identified some part of the cut surface correctly. However, in this same illustration, 64 pupils limited their colouring to only one layer of the cut surface when they might have been expected to have coloured the full extent of the cut surface across the layers. In other words the pupils treated a line as if it was the limit of the cut surface. This pattern was repeated for Hydra where, although 70 pupils identified some part of a cut surface, 66 of them limited their colouring to one layer fo the cut surface. Again, in the Uterus illustration, 57 of the 72 pupils recognised some part of a cut surface but 56 restricted their colouring to only one of the layers of the cut surface represented.

Evidence of the difficulty posed by a change in direction of the plane of a cut surface was less consistent but equally interesting. It should be recalled that the squash ball presented to all pupils had had a quadrant removed thereby exposing a cut surace with a change of direction and that the illustration of the ball likewise showed a change in the direction of the plane of cut. Thus, unlike the two layers within the cut surface, this was a feature which had been presented in the ball tasks. In Alveoli all pupils colouring in the cut surface followed the change in direction correctly; that is, there were no errors associated with the change in direction of the cut surface. By contrast, in the Fish illustration, 41 pupils coloured either the horizontal or the vertical section of the cut surface but not both. In the Spirogyra illustration, which only one pupil coloured completely correctly, four pupils made the error of colouring the cut surface differentially along the length and across the breadth of the cell. As nothing about the objects illustrated would appear to be sufficient to account for these difficulties in colouring, it seems likely that the explanation lies in the means of representation.

Although it was suspected that the non-saggital sectioning portrayed in the Hydra and Uterus illustrations were possible causes of difficulty, none of the colouring patterns suggested that this was so.

While it might be anticipated that the width of the cut surface might cause pupils some difficulty, the results indicate otherwise. The Alveoli illustration, which arguably had the narrowest cut surface, proved to be the easiest to colour correctly.

In Fish, Hydra and Spirogyra some areas of cut surface were not continuous with each other, for example the cut end of the tentacle in Hydra. It was noted that these areas were associated with errors in colouring.

Turning now to the means by which the objects were represented, the illustrations showed an extensive range of conventions used in depicting the object. As none of the illustrations used colour, therefore line and shading alone were available to represent the objects. Kennedy's categories were used as a starting point and the illustrations were analysed in terms of the conventions used. It is clear from a glance that line and shading are not used in the same way in each of the illustrations but systematic analysis revealed that the extent of the variation was very considerable. Line could be used to represent the edge of the object, occluding bound or occluding edge, and also to represent the edge of the cut surface and, again, the layers within the cut surface or a change in its direction. In only one illustration (Fish) was line of the same thickness or heaviness used throughout. An interesting contrast in this respect was shown between Uterus and Alveoli. In Uterus the heaviest line shows the occluding bound of the object, i.e., the line round the outside is the heaviest. However, the feature of significance for the task, the cut surface, is bounded by lighter lines in one part, whereas heavier lines bound the cut surface of another part. Light line was also used to indicate different layers within the cut surfaces of the object. By contrast, in Alveoli, the heaviest line indicates one edge of the cut surface, the feature significant in terms of the task set. In the case of change in direction of the cut surface which caused difficulty in one illustration (Fish) but not others, the feature was represented differently amongst the illustrations. In Fish and Spirogyra a line was present denoting the change of direction but in Alveoli there was no line present. With respect to shading, the illustrations also differ quite considerably. Shading is absent entirely in Fish and Alveoli but present in the other illustrations. It is most evident in Uterus and Blastula. The results of this analysis for each illustration are summarised in Table 2.

The results for Blastula show how this means of analysis provided access to the complexity of the illustrations. Blastula was the shape most like the squash ball. Effectively the object is a hollow sphere with a wall of varying thickness but the illustration uses many conventions to depict this. Line is used to show significant

TABLE 2

PICTORIAL CONVENTIONS USED IN THE ILLUSTRATIONS

| PICTORIAL CONVENTION | ILLUSTRATION | | | | | |
|---|---|---|---|---|---|---|
| | Alveoli (d) | Spirogyra (e) | Hydra (f) | Blastula (g) | Fish (h) | Uterus (i) |
| Heavy line around cut surface | X | | | | | |
| Light line around cut surface | | | | | | X |
| Varied line around cut surface | X | | | X | | X |
| Line left partly "understood" | | | | X | | |
| Line used as occluding edge | X | X | | X | X | X |
| Line used as occluding bound | X | X | X | | X | X |
| Shading used to represent material or colour | | X | X | X | | X |
| Shading used to describe shape | | X | X | X | | X |
| Shading used to represent texture | | | | X | X | X |
| Shading used differentially on cut surface | | X | X | | | X |

X = Pictorial convention used in the illustration features and also texture. Unique amongst the six illustrations line is left broken at some points and must be inferred from a row of barely visible dots. Shading is used to indicate nuclei but not quite as might be expected for modelling their shape.

Although the two aspects of illustrations, the complexity of the object and the means of representation, have been analysed separately it is their interrelationship which is important. For instance, the different layers of the cut surface have been shown to be associated with mistakes but another way of describing these same mistakes is in terms of the variety of means of representation. In the cases where layers are depicted in the cut surface, line is used both to bound the whole of the cut surface and also to discriminate within it. Similarly shading was used to discriminate between layers. In each of the drawings showing layers, some of these layers are differentially shaded.

This detailed attempt to discover the features of illustrations which caused difficulty directs attention to the lack of success the pupils had. Although in uncovering difficulties this is essential, it would be far from correct to suggest that the pupils had no success at all. The pupils showed obvious effort and enthusiasm to colour their work carefully. But besides this the pupils achieved some other successes. All pupils coloured with discrimination some part of the illustration; no pupil went over the edge of the object; pupils generally identified one of the occluding bounds of the cut surface and pupils generally made fine discriminations in colouring between one scoring area of an illustration and another.

## DISCUSSION

The main result of this experiment was that the pupils could not identify the cut surface in five of the six illustrations they were shown. Consequently there is a strong suggestion that it is not safe to assume that 11-year-olds can make enough sense of these illustrations to interpret them correctly; that is to say to interpret them as the teacher might expect. This result has relevance to the three main points in the experiment.

First, the materials were chosen from those available commercially rather than prepared solely for an experiment. It would be unexpected for the pupils to find themselves using the materials in exactly the same way as in the experiment. However, it would be expected to find pupils using the materials in a way which demanded the knowledge which this experiment has demonstrated that they do not have. A teacher might easily expect that a pupil would be able to identify the cut surface in such illustrations. Furthermore, the success of the pupils in identifying some part, but not the complete extent of the cut surface, would not necessarily make the teacher's task easier. This point is not merely a technical one since not only are illustrations widely used in science teaching but the texts from which these particular examples have been drawn are well used (Constable, 1983). The importance of the point is further emphasised when it is recalled that teachers commonly ask pupils to use books in homework tasks. Constable and Brown (1985) note that in one local education authority 15 per cent of science teachers of 11-year-olds report that their pupils used their books for homework. A fifth of this group indicated that this was the only use they made of science texts. It is therefore important that artists, publishers and teachers are all aware of what some pupils might make of these illustrations.

A second aspect of the experiment was the bearing it might have on the question of how pictures are understood and what makes some of them more difficult than others. An important point to emerge from this experiment was that it showed that the difficulties encountered in understanding related to both the complexity of the object being depicted and the means by which it was represented. Further, the difficulties concerned with the means of representation were not simply those one would expect such as size, visibility or number of lines. What seems to be indicated is that a dependence upon familiarity with the pictorial conventions employed is important to understanding the illustration. The analysis indicates that both the particular use of line and shading could lead to the difficulties expressed in the results. The counter-argument that the differences which seemed to be related to difficulty are variations in fidelity to the object represented are more difficult to sustain since these pupils evidently had no knowledge of the appearance of the actual objects illustrated.

This line of investigation has so far produced only a tentative scheme of analysis but further development should lead to testable hypotheses about what makes an illustration more or less accessible to pupils.

The final point of note was the age of the pupils. These 11-year-olds, who had convincingly demonstrated their familiarity with the concept of cut surface, were unable to indicate correctly the extent of cut surface on five out of six illustrations. Although the literature on pictorial perception development does identify certain aspects of picture interpretation which are still incomplete at this age, these tend to be sophisticated or abstract accomplishments such as placing metaphorical meanings on pictures. The import of the present study is that much more basic aspects of picture perception, for example the ability to identify a defined area within a picture, which are normally taken to be well developed by late infancy, may still be in question at 11 years old.

Notwithstanding the clear necessity for further research, it is already possible to draw conclusions from this experiment which can inform the work of producers and users of illustrations. They also inform theoretical studies. Inasmuch as the analysis of the results has, in part, been in terms of the conventions used in the illustrations and that this has proved fruitful, then the notion that such conventions have some part to play in the explanation of the difficulty of picture interpretation is supported. Hence the results do more to support rather than refute the idea of Goodman (1968) that pictures are fundamentally conventional in nature.

From the evidence of the pupils' responses alone, the assumption that they make the same sense of the illustrations as do teachers is distinctly challenged. This has implications for the use of such illustrations in science teaching. The first of these is that there is need for an increased awareness in science teachers of the difficulties which certain illustrations may pose for pupils. Furthermore, it must be recognised that if an illustration is intended to support the introduction of new material then this will be effective only as long as pupils have a sufficient understanding of whatever conventions are used in the illustration. More importantly, if the pupil does not understand the conventions, he may use a dimly understood view of what the object may be, to decipher the conventions.

One way to overcome this may be to teach explicitly about conventions in relation to viewing illustrations. Geography teachers may be a source of information and ideas here. There is some parallel between the conventions used in maps and those employed in illustrations in science, both having a large symbolic content. Geography teachers spend time and attention teaching such conventions and how they do this may be informative. Another perspective is provided by work on scanning pictures. Coles *et al.* (1977) report that adults' scanning strategies are more organised and systematic than those of children. This strategy or lack of it, could contribute to the process by which children lose sight of the task in hand and become distracted by details. Elkind *et al.* (1964) confirm this in a general sense, noting that in viewing complex pictures children's ability to describe the whole increased with age. This then implies that in teaching using illustrations there is a need to emphasise the whole structure of the referrent which is portrayed in the illustration.

It is clear that although picture perception is well researched and that there has been a considerable number of studies concerning pictures and learning, the two fields are not well articulated. It is in this area of intersection that further work is necessary if there is to be progress in understanding of everyday teaching and learning processes. It would be helpful, firstly, to know whether older pupils had as much difficulty with unfamiliar illustrations as younger pupils apparently do and, secondly, to ascertain the picture perception difficulties experienced by pupils considered to be familiar with the objects portrayed. A further valuable step would be to develop the analysis of illustrations into testable hypotheses concerning the difficulties present in illustrations. It is interesting to speculate that one of the outcomes of such work might be the development of a system for awarding a "difficulty score" to illustrations and to ponder on how the intended function of an illustration might influence such scores. Valuable insights to such areas might be gained by exploring the conceptualisations used by illustrators when engaged in producing work for different functions. All these lines of enquiry would lead to the possibility of a more informed use of illustrations as an aid to pupil learning.

ACKNOWLEDGMENTS.—We are grateful to the teachers and pupils who participated in the study and to Sandra Johnson, Assessment of Performance in Science Unit, University of Leeds, for her assistance with the statistical analysis of the results.

Correspondence and requests for reprints should be addressed to Dr. R. M. Campbell, Department of Education, University of York, Heslington, York, YO1 5DD.

## REFERENCES

ARNHEIM, R. (1969). *Visual Thinking*. California: University of California Press.

BEAGLES-ROOS, J., and GREENFIELD, P. M. (1979). Development of structure and strategy in two dimensional pictures. *Dev. Psychol.*, 15, 483-494.

BLAINE, D. (1982). Using pictures to teach concepts to handicapped learners. *Special Education in Canada*, 56, 19-21.

BROWN, R. (1985). The value of pictures in the early years of schooling. *C.O.R.E.*, 9, 2.

CAMPBELL, P. F. (1976). *Literature Review: Research on Children's Comprehension of Picture.* Technical Report 5, Project for Mathematical Development of Children, Florida State University.

COLES, P., SIGMAN, M., and CHESSEL, D. (1977). Scanning strategies of children and adults. In BUTTERWORTH, G. (Ed.), *The Child's Representation of the World.* London: Plenum Press.

CONCANNON, S. J. (1975). Illustrations in books for children: review of research. *The Reading Teacher*, 29, 254-256.

CONSTABLE, H. (1983). Illustrations in science teaching: an investigation into some aspects of their value in learning. MA Dissertation, Department of Education, University of York.

CONSTABLE, H., and BROWN, R. (1986). *Illustrations in First year Science Teaching.* County Durham Research Report II.

DUCHASTEL, P. C. (1978). Pictorial illustration in instructional technology. *Educ. Technol.*, 19, 20-25.

DWYER, F. M. (1972). *A Guide for Improving Visualised Instruction.* Pennsylvania: State College Learning Service.

ELKIND, D., KOEGLER, R., and GO, E. (1964). Studies in perceptual development II: part-whole perception. *Child Dev.*, 35, 81-90.

FLEMING, M. (1979). On pictures in educational research. *Instruct. Sci.*, 8, 235-251.

FRIEDMAN, S. L., and STEPHENSON, M. B. (1975). Developmental changes in the understanding of implied motion in two dimensional pictures. *Child Dev.*, 46, 773-778.

GIBSON, J. J. (1954). A theory of picture perception. *Audio-Visual Communication Review*, 1, 3-23.

GOODMAN, N. (1968). *Languages of Art.* Indianapolis: Bobbs-Merrill.

GOMBRICH, E. H. (1969). *Art and Illusion: a Study in the Psychology of Picture Representation.* Princeton: Princeton University Press.

HALE, G. A. (1971). *Interrelations in Children's Learning of Verbal and Pictorial Paired Associates.* Princeton: Educational Testing Services.

HOLLIDAY, W. G. (1979). Using pictures in the classroom. Paper presented to the National Association for Research in Science

Holliday, W. G., Bruner, L. L., and Donais, G. L. (1977). Differential cognitive and affective responses to flow diagrams in science. *J. Res. Sci. Teaching*, 14, 129-138.

Kennedy, J. M. (1974). *A Psychology of Picture Perception*. San Francisco: Jossey-Bass.

Lippman, M. Z., and Shanahan, M. W. (1973). Pictorial facilitation of paired-associate learning: implications for vocabulary training. *J. educ. Psychol.*, 64, 216-222.

Luczoz, M. A. (1982). A developmental study of recognition and recall of complex pictures. Paper presented to the American Psychological Association.

Nesbit, L. L. (1981). Relationship between eye movements, learning and picture complexity. *Educ. Communic. Technol.*, 29, 109-116.

Nicholson, J. R., and Seddon, G. M. (1977). The influence of secondary depth clues on the understanding by Nigerian schoolboys of spatial relationships in pictures. *Br. J. Psychol.*, 68, 327-333.

Petterson, R. (1982). International review: cultural differences in the perception of image and colour in pictures. *Educ. Communic. Technol. J.*, 30, 43-53.

Samuels, S. J. (1970). Effects of pictures on learning to read, comprehension and attitude. *Rev. educ. Res.*, 40, 397-407.

Serpell, R. (1976). *Culture's Influence on Behaviour*. London: Methuen.

Wittgenstein, F. (1961). *Tractatus Logico-philosophicus*. London: Routledge and Keegan Paul.

(*Manuscript received 28th March*, 1985)

## APPENDIX 1

### Sources of Book Illustrations

| Illustration | Source |
|---|---|
| Alveoli | Mee, A. J., Boyd, P., and Ritchie, D. (1973).<br>*Science for the Seventies* (p. 137).<br>London and Edinburgh: Heinemann Educational Books. |
| Spirogyra | Mackean, D. G. (1973).<br>*Introduction to Biology* (p. 175).<br>London: John Murray. |
| Hydra | Brocklehurst, K. G., and Ward, H. (1970).<br>*A New Biology* (p. 17).<br>London: English University Press. |
| Blastula | Monger, G. (1975).<br>*Revised Nuffield Biology, Text 1* (p. 110). |
| Fish | Mackean, D. G. (1973).<br>*Introduction to Biology* (p. 161).<br>London: John Murray. |
| Uterus | Monger, G. (1975).<br>*Nuffield Biology, Text 1* (p. 98).<br>London: Longman. |

*Br. J. educ. Psychol.*, 58, 103-119, 1988

# CHILDREN FIND FRIENDLY WORDS FRIENDLY TOO: WORDS WITH MANY ORTHOGRAPHIC NEIGHBOURS ARE EASIER TO READ AND SPELL

By VERONICA J. LAXON, VERONIKA COLTHEART
AND CORRIENE KEATING
*(City of London Polytechnic)*

Summary. Forty-nine children (aged 8:4 to 10:4) were grouped according to reading ability and were tested on lexical decision, naming and spelling of words and non-words which differed in orthographic neighbourhood size. Friendly words with many neighbours caused significantly fewer errors on all these tasks. Item analysis demonstrated that these effects were robust when corrected for frequency, regularity, word length and age of acquisition. Good readers showed less reliance on neighbourhood size and more evidence of accurate use of grapheme-phoneme correspondence rules. The results indicate that children find common orthographic sequences easier to read and spell before they have learned to use grapheme-phoneme correspondences consistently.

## INTRODUCTION

There is now considerable evidence to indicate that children rapidly acquire a knowledge of orthographic structure when learning to read. Barron (1981a) points out that orthographic structure incorporates constraints at various levels of abstraction. Morphemic constraints determine different pronunciations of a single digraph such as "th" in "hothead" and in "father" or "ei" in "being" and "rein". Grapheme-to-phoneme correspondence rules determine the different pronunciations of, for example, "a" in "mat" and "mate", or in examples such as "main", "mean", "moan". Phonotactic constraints ensure that sequences difficult to pronounce, e.g., "md", "np" or "tp" do not occur, except when part of separate syllables. Finally, graphotactic constraints are those exemplified in letter sequence rules e.g., in English, words do not begin with "ck", and "q" is invariably followed by "u".

Research on children's reading indicates that their performance in reading tasks is affected by most of the foregoing constraints. The research into the influence of morphemic knowledge in reading tasks has been reviewed by Smith (1984), who concluded that this knowledge increases and develops over several years, from the age of 7. The use of grapheme-phoneme correspondence has also been demonstrated in children as young as 6 in a sentence verification task by Doctor and Coltheart (1980). Other studies have demonstrated their use in non-word (Perfetti and Hogaboam, 1975) and word reading (Seymour and Porpordas, 1980; Backman *et al.*, 1984; Waters *et al.*, 1984). Barron's (1981a) review shows that from the age of 7, children also display a knowledge of phonotactic rules in tasks such as letter detection.

Venezky and Massaro (1977) contrast rule-based descriptions with probabilistic ones, based on letter positional frequencies. A knowledge of these regularities is implied by Henderson and Chard's (1980) findings on lexical decision times for non-words which varied in the extent to which graphotactic rules were embodied. Barron (1980, 1981b) has also found that statistical properties of the letter sequences within words affects children's lexical decision latencies, and that less skilled readers make less use of graphotactic and phonotactic characteristics of words.

Other research which has investigated the (changing) use of orthographic structure in young readers, has adopted a different paradigm. Weber (1970) and Soderbergh (1971) devised indices to compare the graphic similarity of reading errors to the words presented in print. They showed that errors became more graphically similar as reading skill improved. Francis (1984) extended Soderbergh's approach to a larger sample of beginning readers. She found that the graphic similarity of errors to target increased with reading age, and that, initially, this is attributable to the first letter; gradually other letters in the word begin to contribute, with the last letter contributing least. What is more important, this trend to increasing graphic similarity occurred before phonics teaching was given, and at a time when children were unable or unwilling to attempt phonic decoding.

While this evidence of a growth in graphic similarity of errors to target does not indicate what the structure of graphic similarity is, Francis (1984) suggested that it requires a description in terms of positional and sequential probabilities and is not simply explained by grapheme-phoneme conversion rules. Venezky and Massaro (1977) pointed out that use of letter positional frequencies alone can lead to a conflict with phonotactic rules e.g., STHSE, which is unpronounceable, has a higher summed bigram count than SLEVY, which is not. This suggests that a better measure of orthographic structure needs to be devised.

In devising such a measure further evidence needs to be taken into account. Reitsma (1983), while stressing that young children probably use a phonological code as well, presented evidence that orthographic information is very quickly learned and used by 7-year-olds. What is more important, however, he found that the orthographic information is word specific rather than generalised. The children took less time to read non-words, or unfamiliar words on which they had been trained than they did to read non-words which were homophonic with the items on which they had been trained.

Thus any measure of the orthographic structure which children use should involve both word-specific as well as sequential and probabilistic characteristics. Henderson's (1982) description of a lexical pooling model can be used to suggest how such characteristics might be incorporated. According to this model, when a letter string is presented it activates an orthographic neighbourhood of words which differ from the target by one letter in the same position. He distinguishes between *hermits* (e.g., lynx) which have no close neighbours and words with many neighbours; he reports finding substantial effects for lexical hermits in unpublished studies. He suggests that *hermits* are disadvantaged at two stages. At the activation stage it will take longer for activation to spread to remote analogies (for lynx these would be lymph, manx, etc.); at the assembly stage many small segments would be involved, graphemes in the limiting condition, whereas for words with many neighbours, assembly may be from as few as two (larger) segments.

The process then is a probabilistic one, which utilises the specific positions of letters within words; this implies that the orthographic regularities which are influential may include graphemes but are not limited to these, neither are they stored as abstract rules. This definition of neighbourhood, or friendliness then, is not limited to words sharing terminal segments, i.e., the final three letters of four and five letter words as in Glushko's (1981) account. It includes words sharing beginning letters, as well as ends. It is not based on the limited word samples which, as Barron (1981a) points out, have been used to provide letter positional frequencies.

Coltheart *et al.* (1977) have investigated the effect of neighbourhood size in lexical decision tasks with adults. They found that lexical decision for words was unaffected by neighbourhood size. Decision times to reject non-words were significantly longer for non-words with few word neighbours (we term these unfriendly) as compared with those with many (we term these friendly). This suggests that a lexical pooling model does not account for the performance of skilled readers with words already present in their visual word lexicons, but may do so with novel items, not represented in the lexicon, where neighbourhood size may affect performance. Therefore, for children who are still learning to read and for whom many words are still unrepresented (or incompletely represented) in the lexicon, word neighbourhood size, or friendliness, may affect performance. In the present study, Coltheart *et al.*'s (1977) words and non-words, which varied in the size of their neighbourhoods, were presented to children in oral reading (naming) and lexical decision tasks.

Evidence for the use of orthographic structure in spelling is less extensive. However, Sterling (1983) found morphemic effects in the spelling of 12-year-olds. An advantage or regularity in sound-to-spelling correspondence has also been reported for spelling in several studies (Barron, 1980; Frith, 1980; Seymour and Porpodas, 1980). This, together with the fact that spelling errors are often phonetically faithful representations of the word spoken (Bradley and Bryant, 1979; Barron, 1980; Frith, 1980), suggests that children do use phoneme-grapheme correspondence rules sometimes. However, this claim must be qualified in the light of Campbell's (1985) data. She found that only very able, young spellers could use phoneme-grapheme correspondence rules to spell non-words. Poorer spellers used sub-word segments. Ultimately, of course, the competent speller must be capable of retrieving the orthographic representations of entire words since in English there are often several phonologically acceptable representations for a single word: Frith (1985) suggests that this involves the use of full cues and orthographic segments. The good speller becomes progressively more skilful in these ways and thus relies less on the use of phoneme-grapheme conversion, which can produce errors through accurate phonological decoding.

Although a high level of skill in retrieving orthographic information may only be found in the later stages of spelling acquisition, Ehri (1980) has found evidence that even young spellers (aged 7) are quick to acquire orthographic information and use it in spelling. Children had to learn non-words printed in one of two phonetically equivalent ways. For example, they were presented a non-word commencing with "wh" or "w" or with a central "sch" or "st", and later, they had to spell it. Even misspellings tended to preserve the orthography presented earlier, i.e., all misspellings of "wheople" began with "wh" and all those of "weepel" began with "w". Similarly all misspellings of "bistion" contained "st" and all those for "bischun" contained "ch".

In this study, the effect of word friendliness on spelling was also investigated: children were presented the same lists in a spelling task. Children aged 8 to 10 were tested since they could reasonably be expected to represent the ages at which orthographic knowledge has been shown to develop (Barron, 1981a). The children were subdivided according to reading age since the evidence suggests that the effects of friendliness might be expected to differ at different stages of reading attainment.

## METHOD

All the pupils in the second (N = 23) and third year (N = 26) classes of an inner city school were tested. The age range was 8:4 to 10:4 with mean chronological ages of 8:11 and 10:1 for the two classes.

### *Standardised tests and formation of ability groups*

Prior to the experimental sessions the children were given the SPAR standardised group tests of reading and spelling (Young, 1976). There was no difference in the mean reading ages of the two classes (second year x̄ = 8:10, third year x̄ = 8:9); they were therefore treated as one group and subdivided into reader groups on the basis of the raw Spar reading scores. The range of raw scores (14 to 45) was subdivided into three, and these seemed to coincide relatively well with qualitative changes in the reading test. This yielded rather uneven sized groups but no other non-arbitrary method seemed available. Dividing the subjects into three equal-sized groups would have made for a disproportionately wide range of ability in the poorest group. Four children were absent for the Spar Spelling test, and the lexical decision task, three were absent for the naming task and three different children, for the spelling tasks. The sample characteristics are set out in Table 1.

TABLE 1

SAMPLE CHARACTERISTICS MEANS FOR CHRONOLOGICAL AGE AND PERFORMANCE ON STANDARDISED TESTS OF READING AND SPELLING ACCORDING TO READER GROUPS

| | Reading Group | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | T |
| Chronological age | 9·7 | 9:5 | 9:6 | |
| Age equivalent: Reading Age | 10:0 | 8:5 | 7:6 | |
| Number of children | 21 | 20 | 8 | 49 |
| Age equivalent: Spelling Age | 10:8-10:9 | 9:5-9:6 | 8:4 | |
| Number of children | 19 | 18 | 8 | 45 |

*Stimulus materials*

The items used were devised by Coltheart *et al.* (1977): 39 friendly words (mean no. of neighbours = 11·18) and 39 unfriendly words (mean no. of neighbours = 2·00.These were matched pairwise on frequency (Kucera and Francis, 1967) and word length. Thus there were 20 four-letter words and 19 five-letter words in each set, and the frequency range was quite wide (0-686). There were also 40 each of friendly and unfriendly pronounceable non-words. The mean numbers of neighbours were 11·28 and 2·25 respectively and there were 20 each of four- and five-letter items in both lists.

*Lexical decision*

The children were tested individually in a quiet room. The 78 words and 80 non-words were randomised and presented singly on the video display of a BBC micro-computer. The presentation was self-paced, i.e., the children were instructed to press the space bar to make a new item appear. They were instructed to make a "no" response if the item was not a word and a "yes" response if it was a word. A voice key triggered by the onset of vocalisation caused the timer to stop. Reaction times and the correctness of the response were recorded. An initial practice session preceded the main experiment to ensure that the child understood the task.

*Naming*

The children were then tested, individually, on a separate occasion. As before, the subject-paced items were presented in random order on a BBC micro-computer but now the children had to read each one aloud. The children were told that half the items were nonsense words, and asked to try to read them as best they could.

*Spelling*

The items were given as a group test a month later in a single order, words being presented before non-words. Where words were ambiguous or homophonic these were put into a sentence in order to make the meaning clear. For non-words the children were again instructed that these were nonsense words and that they should try to write them how they thought they should be spelled.

The possibility that order effects might have influenced the results was considered and so these items were administered to a second sample of 24 children in another inner city school. Two different random orders were constructed and words and non-words were interspersed. Error scores for each word were calculated for both samples and the resulting correlation was 0·922. This seemed to indicate that the blocking of similar items for the main sample did not affect the spelling performance on individual items in any special way.

## RESULTS

### I. PERFORMANCE ANALYSIS

### *Inter-correlations between standardised tests and experimental tasks*

The inter-correlations between standardised tests and experimental items are set out in Table 2 which shows that all the correlations were significant and high.

TABLE 2

INTER-CORRELATIONS BETWEEN STANDARDISED TEST (RAW) SCORES AND PERFORMANCE ON EXPERIMENTAL TASKS

| | *Experimental Tasks* | | | |
|---|---|---|---|---|
| | Reading | | Spelling | |
| | Words | Non-words | Words | Non-words |
| *Standardised tests* | | | | |
| Spar reading | 0·81 | 0·77 | 0·82 | 0·64 |
| Spar spelling | 0·84 | 0·90 | 0·93 | 0·77 |
| *Experimental tasks* | | | | |
| Word reading | | 0·88 | 0·86 | 0·74 |
| Non-word reading | | | 0·92 | |
| Non-word spelling | | | 0·85 | |

### *Analysis of accuracy on experimental tasks*

Analyses of variance were carried out using reader group as a between-subjects factor, and friendliness as a within-subjects factor for each of the three experimental tasks. Separate analyses were performed for non-words and words and for each of the experimental tasks, lexical decision, naming and spelling.

### *Lexical decision*

The mean correct lexical decisions for words (yes decisions) and non-words (no decisions) are set out in Table 3. The reader groups differed significantly in accuracy (F(2, 46) = 15·47, MSe = 22·37, P < 0·001). Planned comparisons indicated that Group 1 performed better than Group 2 (t(46) = 3·92, P < 0·001) who performed better than Group 3 (t(46) = 2·15, P < 0·05).

Accuracy for unfriendly words was significantly lower than for friendly words (F(1, 46) = 86·55, MSe = 3·51, P < 0·001). The interaction of friendliness with reader group was also significant (F(2, 46) = 6·60, P < 0·003). Planned comparisons indicate that all groups were more accurate with friendly words

TABLE 3

MEAN PER CENT CORRECT SCORES AND STANDARD DEVIATIONS FOR LEXICAL DECISIONS ON FRIENDLY AND UNFRIENDLY WORDS AND NON-WORDS ACCORDING TO READER GROUP (SDs IN PARENTHESES)

| *Words*<br>*Reader Group* | *N* | Friendly | Unfriendly |
|---|---|---|---|
| 1 | 21 | x̄ 88·9 | 83·8 |
| | | (19·54) | (8·67) |
| 2 | 20 | x̄ 80·7 | 71·0 |
| | | (19·95) | (8·95) |
| 3 | 8 | x̄ 75·6 | 59·4 |
| | | (14·43) | (13·97) |
| *Non-words*<br>*Reader Group* | *N* | | |
| 1 | 21 | x̄ 71·4 | 82·8 |
| | | (18·53) | (12·55) |
| 2 | 20 | x̄ 71·1 | 78·5 |
| | | (18·38) | (16·10) |
| 3 | 8 | x̄ 48·1 | 56·9 |
| | | (12·53) | (14·38) |

(t(46)=3·46, P < 0·001 for Group 1, t(46)=3·75, P < 0·001 for Group 2 and t(46)=6·33, P < 0·001 for Group 3). The significant interaction is attributable to the fact that Group 1 was significantly less affected by friendliness than Group 2 (t(46)=2·11, P < 0·04) whereas Group 2 and 3 showed comparable, large effects of friendliness. The stimulus analysis confirmed the significant difference between the reader groups (F(2, 152) = 40·87, MSe = 170·05, P < 0·001), friendliness (F(1, 76) = 6·05, MSe = 958·43, P < 0·02) and the significant interaction (F(2, 152) = 3·11, MSe = 170·05, P < 0·05).

The reader groups differed in accuracy on non-word lexical decision (F(2, 46) = 7·77, MSe = 77·03, P < 0·001). Groups 1 and 2 did not differ significantly in accuracy, whereas Group 2 were more accurate than Group 3 (t(46) = 3·44, P < 0·001). Unfriendly non-words were significantly easier to reject than friendly non-words (F(1, 46) = 30·34, MSe = 8·92, P < 0·001). There was no significant interaction between friendliness and reader group. Response latencies were not analysed in view of the very high error rates and the large effect of friendliness on accuracy.

The stimulus analysis confirmed the significant reader group effect (F(2, 156) = 105·14, P < 0·001) and friendliness (F(1, 78) = 16·57, MSe = 366·67, P < 0·001). The interaction was not significant.

### *Naming and spelling of words*

The mean correct scores for naming and spelling of words are set out in Table 4. The reader groups differed significantly in accuracy in reading the words aloud (F(2, 43)=38·19, MSe=52·48, P < 0·001). Group 1 readers were more accurate than Group 2 (t(43)=2·15, P < 0·05) who performed better than Group 3 (t(43)=6·88, P < 0·001).

TABLE 4

Mean Per Cent Correct Scores and Standard Deviations for Naming and Spelling of Friendly and Unfriendly Words According to Reader Group (SDs in Parentheses)

| *Naming*<br>*Reader Group* | *N* | Friendly | Unfriendly |
|---|---|---|---|
| 1 | 20 | x̄ 95·6 | 91·3 |
| | | (4·86) | (6·14) |
| 2 | 18 | x̄ 89·9 | 78·6 |
| | | (8·72) | (9·47) |
| 3 | 8 | x̄ 53·2 | 38·5 |
| | | (28·58) | (29·90) |
| *Spelling*<br>*Reader Group* | *N* | | |
| 1 | 21 | x̄ 90·1 | 70·1 |
| | | (7·49) | (14·24) |
| 2 | 17 | x̄ 79·8 | 53·1 |
| | | (19·38) | (21·49) |
| 3 | 8 | x̄ 38·46 | 14·1 |
| | | (19·53) | (11·22) |

Unfriendly words were significantly harder to *read* than friendly words (F(1, 43) = 76·44, MSe = 3·98, P < 0·001). The interaction of friendliness with reader group was significant (F(2, 43) = 7·51, P < 0·005). The results of planned comparisons (t(43) = 2·69, P < 0·01 for Group 1, t(43) = 6·60, P < 0·001 for Group 2 and t(43) = 5·76, P < 0·001 for Group 3) indicate that the effect of friendliness held for all three groups. The significant interaction is, again, attributable to the fact that Group 1 was significantly less affected by friendliness than Group 2 (t(43) = 2·93, P < 0·005), and that Groups 2 and 3 showed comparable large effects. The stimulus analysis confirmed the significant effect of reader group (F(2, 152) = 197·04, MSe = 250·34, P < 0·001) and friendliness (F(1, 76) = 5·81, MSe = 806·15, P < 0·02). The interaction, however, was not significant.

The reader groups differed significantly in spelling accuracy (F(2, 43) = 37·03, MSe = 68·97, P < 0·001). Again, Reader Group 1 were more accurate than Reader Group 2 (t(43) = 2·78, P < 0·01) whose performance was more accurate than that of Group 2 (t(43) = 6·22, P < 0·01). Unfriendly words were significantly harder to *spell* than friendly words (F(1, 43) = 179·06, MSe = 9·27, P < 0·001). The interaction of friendliness and reader group was not significant.

The stimulus analysis confirmed the significant effect of reader group (F(2, 152) = 333·26, MSe = 177·86, P < 0·001) and friendliness (F(1, 76) = 25·00, MSe = 1125·64, P < 0·001). The interaction was not significant.

### *Scoring of naming and spelling of non-words.*

The scoring of performance on non-words is not straightforward. For example, several items lend themselves to analogical decoding in reading, e.g., *mough* and *gight*. Others could be regarded as being covered by special environment rules (Wijk, 1966), e.g., *mook, shold, jild, wrolk*. Yet others involved minor correspondences, or could be read analogously with an exception word, e.g., *bave, sare, chone.*

Because the possible range of acceptable grapheme-phoneme correspondences are not symmetrical for reading and spelling, rather different considerations arose in scoring the spelling items. In order to discover the extent to which children use abstract grapheme-phoneme rules, naming responses were scored correct only when the pronunciation conformed to Wijk's (1966) classification of the regular form. In the case of spelling, various alternative forms were taken to be acceptable for some non-words e.g. |kluːn| → "cloon", "klune", "kloon". The mean correct scores for naming and spelling of non-words is set out in Table 5. The experimenter who read out the items occasionally diverged from the strictly "regular" decoding, pronouncing, for e.g., *mough* to rhyme with *though*, therefore "mow", "moae" and "mowe" were scored "correct".

TABLE 5

MEAN PER CENT CORRECT SCORES AND STANDARD DEVIATIONS FOR NAMING AND SPELLING OF FRIENDLY AND UNFRIENDLY NON-WORDS ACCORDING TO READER GROUP (SDS IN PARENTHESES)

| *Naming* *Reader Group* | *N* | Friendly | Unfriendly |
|---|---|---|---|
| 1 | 20 | 73·4 | 67·6 |
| | | (13·68) | (16·97) |
| 2 | 18 | 61·6 | 54·6 |
| | | (22·21) | (16·71) |
| 3 | 8 | 25·0 | 20·9 |
| | | (28·19) | (20·95) |
| *Spelling* *Reader Group* | *N* | | |
| 1 | 21 | x̄ 65·6 | 55·1 |
| | | (19·56) | (22·30) |
| 2 | 17 | x̄ 55·1 | 35·4 |
| | | (28·10) | (21·20) |
| 3 | 8 | x̄ 26·9 | 16·6 |
| | | (28·56) | (15·69) |

The reader groups differed significantly in non-word reading accuracy (F(2, 43) = 19·38, MSe = 106·73, P < 0·001). Group 1 were more accurate than Group 2 (t(43) = 2·09, P < 0·05) and Group 2 were more accurate than Group 3 (t(43) = 4·53, P < 0·001). Unfriendly non-words were significantly less accurately named than friendly non-words (F(2, 43) = 12·59, MSe = 7·87, P < 0·001). The interaction with reader group was not significant. The stimulus analysis confirmed the significant effect of reader group (F(2, 156) = 294·82, MSe = 163·59, P < 0·001). Friendliness just failed to reach significance (F(1, 78) = 2·76, MSe = 8·28, P < 0·10). The interaction was not significant.

The reader groups differed in accuracy in spelling non-words (F(2, 43) = 9·09, MSe = 155·97, P < 0·001). Performance of Group 1 was superior to that of Group 2 (t(43) = 2·09, P < 0·05) whose performance exceeded that of Group 3 (t(43) = 2·49, P < 0·05). Unfriendly non-words were significantly harder to spell than friendly non-words. (F(1, 43) = 51·32, MSe = 11·05, P < 0·001). The interaction between friendliness and reader group was significant (F(2, 43) = 3·34, P < 0·05).

The results of planned comparisons indicate that unfriendly non-words were significantly harder than friendly for all groups ($t(43) = 4\cdot08$, $P < 0\cdot001$ for Group 1, $t(43) = 6\cdot91$, $P < 0\cdot001$ for Group 2 and $t(43) = 2\cdot48$, $P < 0\cdot05$ for Group 3), and the interaction is attributable to the fact that Group 1 were significantly less affected by friendliness than Group 2 ($t(43) = 2\cdot41$, $P < 0\cdot02$); whereas Group 2 and 3 were similarly affected by friendliness. The stimulus analysis confirmed the significant effect of reader group ($F(2, 156) = 231\cdot18$, $MSe = 113\cdot58$, $P < 0\cdot001$) and friendliness ($F(1, 78) = 14\cdot26$, $MSe = 869\cdot75$, $P < 0\cdot001$). The interaction was also significant ($F(2, 156) = 5\cdot54$, $MSe = 113\cdot58$, $P < 0\cdot005$).

## II. ITEM ANALYSIS OF WORD NAMING AND SPELLING

Although the analysis over subjects yielded highly significant effects of friendliness, the robustness of these effects was further examined in a detailed stimulus item analysis which included other measures known to affect difficulty. Difficulty values were calculated for each word by counting the number of errors made on each word for both spelling and naming by the children. The measures used, in addition to friendliness, were:

*Frequency.* The words used had been matched on the Kučera Francis (1967) word frequencies. However, this estimate may not be the most appropriate for English children. Therefore, frequencies were obtained for each word from Hofland and Johansson's (1982) British printed word count and Carroll *et al.*'s (1971) word frequencies in American children's reading matter. Two measures were explored from the latter, the frequency for Grade 3 (since the children used in our study were close in age to those in Grade 3), and the overall frequency. The latter (AHF) was used since it accounted for most variance. Log values were used in view of findings by Theios and Muise (1977).

*Regularity.* A strict dichotomous definition of regularity in terms of use of a major correspondence rule was devised (Venezky, 1970) and special environment rules (Wijk, 1966) were disregarded. Two other classifications were also devised which took into account degree and nature of irregularity in view of the findings by Backman *et al.* (1984) and Waters *et al.* (1984). These also predicted substantial amounts of variance, suggesting that further investigation of these would be appropriate. However, they were not noticeably better than the dichotomous definition, when used with the other predictors; and, since their use resulted in sub-cells of markedly unequal size with this particular word set, they were not used in this analysis.

*Word length* (WL). Although the lists were matched on word length, word length was also used in view of findings such as those of Seymour and MacGregor (1984).

*Age of acquisition* (AOA). In view of a recent finding (Coltheart *et al.*, 1988) of the pervasive effects of Age of Acquisition, it was thought important to obtain effects corrected for this variable; this was especially important in view of the fact that the lists had not been devised specially for children. Accordingly, Age of Acquisition ratings were elicited from 40 undergraduate students at the City of London Polytechnic. These ratings were elicited according to the method used by Gilhooly and Logie (1980). Mean Age of Acquisition ratings were used as a predictor.

A series of logistic models was fitted using GLIM (Baker and Nelder, 1978), a procedure which is analogous to regression, but which uses a binomial error, as appropriate for proportional data, and the transformation ($\log_e \left|\frac{p}{1-p}\right|$). It produces estimates for the predictors which are corrected for all other terms in the model. The estimates of parameters and the semi-partial r's for the fitted models are set out in Table 6.

TABLE 6

ESTIMATES OF THE PARAMETERS AND PARTIAL R'S FOR PREDICTORS OF NAMING AND SPELLING WORD DIFFICULTY

| *Naming Predictor* | b | $^{t}(72)$ | P | semi-partial r |
|---|---|---|---|---|
| GM | −3·00 | | | |
| Friendliness | 0·30 | 3·15 | < 0·01 | 0·13 |
| Regularity | −0·45 | 4·44 | < 0·001 | 0·17 |
| AOA | 0·56 | 9·33 | < 0·001 | 0·38 |
| AH frequency | −0·13 | 4·81 | < 0·001 | 0·19 |
| Word length | 0·47 | 4·98 | < 0·001 | 0·20 |
| *Spelling Predictor* | | | | |
| GM | −2·58 | | | |
| Friendliness | 0·94 | 12·11 | < 0·001 | 0·42 |
| Regularity | −0·56 | 6·75 | < 0·001 | 0·23 |
| AOA | 0·52 | 9·87 | < 0·001 | 0·35 |
| AH frequency | 0·60 | 2·69 | < 0·01 | 0·09 |
| Word length | 0·66 | 8·45 | < 0·001 | 0·29 |

The full models explain a significant and substantial proportion of variance, but the fit was better for spelling ($R^2 = 0{\cdot}624$) than for naming ($R^2 = 0{\cdot}510$). It can be seen that all the predictors explain a significant amount of variance; and that friendliness is significant even when corrected for the other four measures. It should be noted that the semi-partial correlations yielded exclude variation attributable to the joint prediction of two or more variables; therefore where these are substantial, these estimates given are usually an underestimate of their overall predictive power. In particular, the exclusion of AOA results in an increased semi-partial correlation coefficient for AHF and vice versa; this can be explained by a correlation of 0·56 between these two measures. The correlation between friendliness and AOA of 0·21 must also indicate that the semi-partial r for each of these variables is an underestimate. Therefore, caution is indicated when interpreting the results of such an analysis, in terms of extent of an effect (or absence of an effect), since this is affected by the nature of the variables in the model and their inter-relationships.

## III. QUALITATIVE ERROR ANALYSIS

### *General classification of all items*

The items used in the lists were too heterogeneous to allow a very detailed error analysis; therefore errors were subclassified into omissions, words, neologisms and items difficult to classify (e.g., single letters) were termed unclassified. These are reported in terms of percentages of errors and need to be regarded with some caution since they are not calculated on an average per child and could therefore reflect strategies of individual children within the sample. A perusal of the children's protocols, however, suggested that this applied only to omissions and unclassifieds and that these were produced by a few of the less competent children. Unclassified responses constituted 1-2 per cent of all errors; omissions varied as a function of overall difficulty being as low as 5 per cent on non-word spelling of friendly items and 60 per cent on spelling unfriendly word items.

When the above categories were excluded, the proportion of lexicalisations were consistently higher for naming tasks (57 per cent) as compared with spelling (29 per cent), and consistently higher for word items (48 per cent) than non-word items (38 per cent) and consistently higher for friendly items (50 per cent) as compared with unfriendly items (36 per cent). This seems to indicate that the harder the task overall as defined in terms of the means given in Tables 4 and 5, the fewer the lexicalisations. Disparities between the task requirements meant that further word and non-word comparisons were inappropriate.

### *Further analysis of word items*

It should be noted that the words in the lists differed in the extent to which they lent themselves to some of the different categories of errors; as such, the analyses presented below should only be regarded as suggestive. Also, most errors shared the first letter, at least, with the target item. Where this was not the case, they were usually the product of a less able child or where this was not so, probably attributable to mishearing (e.g., *club* → "plug") or misarticulation due to the cockney accent (e.g., *thorn* → "fawn", in naming and *axle* → "axe saw", *jury* → "jewellery" in spelling), though for spelling the extra assumption of the use of an articulatory loop (Sterling, 1983) is needed.

The lexicalisation errors were further subclassified according to their similarity to the stimulus word. Dissimilar items were those sharing less than 50 per cent of the letters. The rest were termed similar and were subdivided into order errors (e.g., *barn* → "bran"), neighbours, or sharing all letters, bar one, in the appropriate place (e.g., *sung* → "song") and "others", (e.g., *bold* → "blond", *scene* → "scorn"). The latter category also included a few errors which could be attributable to the cockney accent in both spelling and naming (*spine* → "spain", *bill* → "build"). A further category was appropriate only for spelling, i.e., homophonic alternatives; five friendly words and six unfriendly words had common homophones. These are shown as percentages of total lexicalisation errors in Table 7.

The neologistic errors were also subclassified:

(a) Order errors (e.g., *once* → "ocne").

(b) Phonic attempts. In spelling these yielded a homophonic non-word (e.g., *rare* → "rair"), while in naming these yielded stress errors (e.g., "*be*low", "maj*or*") or regularisations (e.g. *scene* as though spelled "skene").

(c) Use of alternative rules. In spelling this will yield a homophone if a lesser correspondence rule is used (e.g., *gnaw* → "noor", presumably analogous with *door*). In naming this will produce irregularisation (e.g., *gown* to rhyme with *bone*).

(d) Rule violation. These were violations of the long/short vowel rule (e.g., *fated* → "fatted"); and for spelling only, an additional subclassification was appropriate, violation of "other rule", i.e., g/j, i/y, s/c and e/y in final position.

TABLE 7

LEXICALISATIONS IN NAMING AND SPELLING OF WORDS SUBCLASSIFIED ACCORDING TO SIMILARITY, FRIENDLINESS AND PERCENTAGE OF THE LEXICALISATION ERRORS

| | *Percentage of error type* | | | | |
|---|---|---|---|---|---|
| | order<br>% | Similar neighbour<br>% | other<br>% | Dissimilar<br>% | Homophone<br>% |
| *Naming* | | | | | |
| Friendly | 5 | 23 | 59 | 12 | NA |
| Unfriendly | 1·2 | 13 | 78 | 10 | NA |
| *Spelling* | | | | | |
| Friendly | 23 | 10 | 62 | 4 | 2 |
| Unfriendly | 5 | 1 | 49 | 17 | 30 |

(e) Other. This category included a variety of vowel and/or consonant substitutions; in spelling, liquids in the penultimate position were often omitted. A few were attributable to the cockney accent (e.g. *gown* → "garn"). About a quarter of the spelling items in this category clearly revealed specific word knowledge, though this can usually be unequivocally established only for the orthographically irregular or unusual words (e.g., *eight* → "eighit", *yacht* → "yaght", "yaucht", *itch* → "eaitch").

These are shown as percentages of neologistic errors in Table 8.

TABLE 8

NEOLOGISTIC ERRORS IN NAMING AND SPELLING OF WORDS SUBCATEGORISED ACCORDING TO TYPE OF ERROR, FRIENDLINESS, AND PERCENTAGES OF TOTAL NEOLOGISTIC ERRORS

| | | | | *Rule violation* | | |
|---|---|---|---|---|---|---|
| | order<br>% | Phonic attempts<br>% | Alternative rules<br>% | Long/short vowel<br>% | Other rule<br>% | Other<br>% |
| *Naming* | | | | | | |
| Friendly | 0 | 4 | 30 | 27 | NA | 40 |
| Unfriendly | 6 | 32 | 24 | 1 | NA | 36 |
| *Spelling* | | | | | | |
| Friendly | 12 | 23 | 2 | 17 | 3 | 65 |
| Unfriendly | 12 | 32 | 3 | 6 | 9 | 71 |

### *Naming and spelling of non-words*

The naming and spelling of non-words were included in the hope that this would allow some conclusions regarding the children's use of grapheme-phoneme correspondence rules. The constraints on the selection of non-word items, i.e., that they should vary according to the number of word neighbours, involve four or five letters, and the fact that they were derived from one such neighbour by altering one letter, meant that the lists were very heterogeneous with respect to the rules embodied.

There were, however, some clear indications that the children were not affected in any simple fashion by the complexity of rules embodied in the various items, as has been suggested by Marsh *et al.* (1980, 1981), Guthrie and Siefert (1977) and Venezky and Johnson (1975). The constraints on the selection of items meant that many unfriendly items involved simple vowels. These are less complex than long vowels in terms of graphemic representation according to Venezky and Johnson (1975) in that the latter require vowel digraphs or a final e; yet these were significantly harder than the friendly items where the constraints produced large numbers of items with long vowels.

Also within the friendly items involving long vowels and, therefore, comparable rule complexity, there were marked divergencies as to whether or not these were adequately decoded or encoded. In naming, for example, *hage* had 27 "errors", *rame* had 11 and *yole* 8; for spelling the "errors" were 12, 13 and 29 respectively. This indicates little consistency within or between tasks.

It was also evident that the children were not systematically applying rules irrespective of friendliness value. It was possible to select a small subset of items from the two lists which were comparable in terms of length, complexity and type of rule embodied and to compare the "errors" for naming and spelling. These are set out in Table 9. It can be seen that, in general, unfriendly items are more difficult than comparable friendly items, with the exception of *glime* cf. *blane* in spelling. Also the correspondence between spelling and naming difficulty is not consistent when the number and type of phonemes are the same.

TABLE 9

ERRORS FOR NAMING AND SPELLING OF COMPARABLE UNFRIENDLY AND FRIENDLY NON-WORDS

| *Friendly* | *Naming* | *Spelling* | *Unfriendly* | *Naming* | *Spelling* |
|---|---|---|---|---|---|
| shart | 9 | 19 | tharn | 31 | 35 |
| tink | 7 | 15 | vond | 11 | 26 |
| hage | 27 | 12 | wege | 38 | 38 |
| blane | 15 | 20 | glime | 23 | 19 |
| glave | 14 | 17 | clune | 32 | 29 |
| flad | 10 | 10 | blid | 22 | 21 |
| flack | 7 | 8 | crich | 13 | 24 |

There were some instances of use of special environment rules in the spelling of non-words (e.g., *yole* → "yowl", "yoll"). In naming this was the preferred response with some items which either had a large number of, or several high frequency neighbours. *Mook, wull, shold* and *baft* were typically read according to special environment rules, i.e., analogously with "book", "pull", "old" and "daft". *Jild* and *wrolk*, however, were read with short vowels indicating use of grapheme-phoneme correspondences. There were also several examples where one high frequency exception word or omission of the final e affected pronunciation — the median response to *chone* was to read it analogously with "gone" (? or "shone"); two children read it analogously with "done".

There was evidence in naming that some "errors" of failure to lengthen the vowel were produced where another lexical item exists as suggested by Saffran (1982) (i.e., *hage, rame, mide, tane, stabe*), but this was not consistent in terms of number of errors and frequency of the lexical item in question, and was rare in spelling, in contrast with Read's (1971) data on younger children. Cockney articulation errors seemed to produce some deviant responses in both naming and spelling (e.g., *baft* → "bathed", *bave* → "bathe", *tharn* → "farn"). However, more systematic analyses of this sort were precluded by the differences in extent to which items lent themselves to these sorts of errors. But further research is indicated.

## DISCUSSION

The results indicate that orthographic structure is an important variable in these children's reading in that unfriendly items with few neighbours proved significantly more difficult than those with many, in lexical decision, naming of words and non-words; and this is consistent with the findings of Henderson and Chard (1980) and Barron (1981a, 1981b). The results for non-words are consistent with Coltheart *et al.*'s (1977) adult data. While the effects for words are not, good readers are significantly less affected by friendliness than the other groups, suggesting a possible developmental trend.

The results also indicate that grapheme-phoneme conversion when defined as an abstract system of rules cannot straightforwardly explain the performance of these children. There was a very large percentage of lexicalisation errors for both words (71 per cent) and non-words (57 per cent): this indicates a use of a visual code and/or partial cues (Frith, 1980). The significant effect of friendliness suggests some sort of abstraction is taking place and that children are learning about the orthographic properties of words. However, what they are learning is not simply captured by, or restricted to, abstract grapheme-phoneme correspondence rules. The significant friendliness effect and the lowered overall levels of performance in the reading of non-words suggests that if grapheme-phoneme correspondences are used only those frequently embodied in words are accurately used and those may not be the simplest letter to sound correspondences. The friendly non-words, although balanced for "regularity" with the unfriendly, contained far more items involving long vowels which have been suggested to involve more sophisticated rules (Marsh *et al.*, 1981; Venezky and Johnson 1975; Guthrie and Siefert 1977), i.e., in that they require vowel digraphs or a vowel plus final e, whereas the majority of unfriendly items involved short vowels only. Also words involving rules of equal status differed markedly in difficulty.

The facts that the better readers were significantly less affected by friendliness in lexical decision and word naming, and were able to read non-words with some competence, suggest that this group have developed a fairly well specified lexicon. It also suggests that they have learned to abstract features to the extent that they appear to be using something that looks more like an abstract grapheme-phoneme correspondence system than the processes used by the less able readers. This is, however, by no means complete since their reading of even simple CVCC or CCVC words was not perfect (more than half of them made errors on *crem* for example).

The picture then that emerges from this data is of children gradually learning more about orthographic structure and demonstrating this by increasing abstractions involving some idea of invariance. This may evolve into grapheme phoneme correspondence rules at later levels of sophistication. This account is more consistent with those of Francis (1984) and Campbell (1985) than with the qualitative age stage model suggested by Marsh *et al.* (1981). It is consistent with Henderson's (1982) account of a lexical pooling model, to the extent that it suggests how a lexical pooling model may evolve into a dual-route model.

An alternative explanation of the orthographic neighbourhood effect may be provided by Hinton *et al.*'s (1986) model of distributed representations. This suggests the lexicon is organised as sets of similar words, of the same length, sharing all letters, bar one, in the same positions. Thus, individual words are represented in all the sets which involve their neighbours. Furthermore, Hinton demonstrated that loss and recovery of information after specific training also generalises in a distributed representation system. Coltheart (1985) has reported evidence of this generalisation of recovery of word reading by a surface dyslexic. If a distributed representation characterises the adult visual word lexicon, then children presumably have to acquire such a system. A word possessing many neighbours may perhaps be more readily acquired since it will be represented in many word sets and can more readily be incorporated in the existing system of word sets. Accurate naming and lexical retrieval then could be attributed to the extent of the multiple representation: though caution is indicated in extending this explanation to skilled readers in view of the Coltheart *et al.* (1977) data.

The performance of the children on spelling, given that they show less overall competence, is remarkably similar and lends little support to the suggestion that these children are using markedly different strategies for reading and spelling. Friendly words were spelled as well as unfriendly words were read and friendly non-words were spelled *better* than unfriendly non-words were read. This, taken together with the fact that the friendliness effects were more pronounced for spelling, suggests that children are using orthographic structure in spelling as suggested by Ehri (1980) and that they do this before fully-fledged abstract grapheme-phoneme correspondence rules are available as suggested by Campbell (1985).

The data also suggest that these children may be using partial orthographic cues for both reading and spelling (Frith, 1980) if we take these to mean some partial abstraction from a word incompletely represented in the lexicon. As she points out, this strategy is fairly effective for reading but accurate spelling requires the use of full orthographic representation. Where these are not available, children may be forced to rely on what abstractions they do know in order to spell; and these abstractions may vary in degree and depend on observed and frequent regularities in the orthographic structure of the words to which they are exposed. This would result, as observed here, in the increased significance of friendliness and regularity and lower levels of accuracy and lexicalisation for spelling as compared with reading.

Correspondence and requests for reprints should be sent to Dr. V. J. Laxon, Psychology Department, City of London Polytechnic, London, E1 7NT.

## REFERENCES

Backman, J., Bruck, M., Hebert, M., and Seidenberg, M. S. (1984). Acquisition and use of spelling-sound correspondences in reading. *J. exp. Child Psychol.*, 38, 114-133.

Baker, R. J., and Nelder, J. A. (1978). *The Glim Release 3. Generalised Linear Interactive Modelling.* Oxford: Numerical Algorithms Group.

Barron, R. W. (1980). Visual-orthographic and phonological strategies in reading and spelling. In Frith, U. (Ed.), *Cognitive Processes in Spelling.* London: Academic Press.

Barron, R. W. (1981a). Reading skills and reading strategies. In Lesgold, A. M., and Perfetti, C. (Eds.), *Interactive Processes in Reading.* Hillsdale. NJ: Lawrence Erlbaum Associates.

Barron, R. W. (1981b). Development of visual word recognition: a review. In McKinnon, G. E., and Waller, T. G. (Eds.), *Reading Research: Theory and Advances.* New York: Academic Press.

Bradley, L., and Bryant, P. E. (1979). Independence of reading and spelling in backward and normal readers. *Dev. Med. Child Neurol.*, 21, 504-514.

Campbell, R. (1985). When children write nonwords to dictation. *J. exp. Child Psychol.*, 40, 152-180.

Carroll, J. B., Davies, P., and Richman, B. (1971). *The American Heritage Word Frequency Book.* Boston: Houghton Mifflin.

Coltheart, M., Davelaar, E., Jonassen, J. T., and Besner, D. (1977). Access to the internal lexicon. In Dornic, S. (Ed.), *Attention and Performance. Vol. 6.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Coltheart, V., Laxon, V. J., and Keating, C. (1988). Effects of word imagery and age of acquisition on children's reading. *Br. J. Psychol.*

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In Posner, M. I., and Marin, O. S. M. (Eds.), *Attention and Performance XI.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Doctor, E. A., and Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory and Cognition*, 8, 195-209.

Ehri, L. C. (1980). The development of orthographic images. In Frith, U. (Ed.), *Cognitive Processes in Spelling.* London: Academic Press.

Francis, H. (1984). Children's knowledge of orthography in learning to read. *Br. J. educ. Psychol.*, 54, 8-23.

Frith, U. (1980). Unexpected spelling problems. In Frith, U. (Ed.), *Cognitive Processes in Spelling.* London: Academic Press.

Frith, U. (1985). The usefulness of the concept of unexpected reading failure. Comments on reading retardation revisited. *Br. J. dev. Psychol.*, 8, 15-17.

Gilhooly, K. J., and Logie, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1,944 words. *Behav. Res. Meth. Instrum.*, 12, 395-427.

Glushko, R. J. (1981). Principles for pronouncing print: the psychology of phonology. In Lesgold, A. M., and Perfetti, C. (Eds.), *Interactive Processes in Reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Guthrie, J. T., and Siefert, M. (1977). Letter-sound complexity in learning to identify words. *J. educ. Psychol.*, 69, 686-696.

Henderson, L. (1982). *Orthography and Word Recognition.* London: Academic Press.

Henderson, L., and Chard, J. (1980). Knowledge of orthographic structure. In Frith, U. (Ed.), *Cognitive Processes in Spelling.* London: Academic Press.

Hinton, G. E., McClelland, J. L., and Rumelhart, D. E. (1986). Distributed representations. In Rumelhart, D. E., McClelland, J. L., and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition Volume 1: Foundations.* Cambridge, Mass: MIT Press.

Hofland, K., and Johansson, S. (1982). *Word Frequencies in British and American English.* Bergen: Norwegian Computing Centre for the Humanities.

Kučera, H., and Francis, W. N. (1967). *Computational Analysis of Present-day American English.* Providence, RI: University Press.

Marsh, G., Friedman, M., Welch, V., and Desberg, P. (1980). The development of strategies in spelling. In Frith, U. (Ed.), *Cognitive Processes in Spelling.* London: Academic Press.

Marsh, G., Friedman, M., Welch, V., and Desberg, P. (1981). A cognitive-developmental theory of reading acquisition. In McKinnon, G. E., and Waller, T. G. (Eds.), *Reading Research: Advances in Theory and Practice.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Perfetti, C. A., and Hogaboam, T. (1975). The relationship between single word decoding and reading comprehension skill. *J. educ. Psychol.*, 67, 461-469.

Read, C. (1971). Pre-school children's knowledge of English phonology. *Harv. Educ. Rev.*, 41, 1-34.

Reitsma, P. (1983). Printed word reading in beginning readers. *J. exp. Child Psychol.*, 36, 321-339.

Saffran, E. (1982). Lexicalisation v "rule" application in a surface dyslexic. Paper presented to the Conference on Surface Dyslexia and Surface Dysgraphia, Oxford.

Seymour, P. H. K., and MacGregor, C. M. (1984). Developmental dyslexia: a cognitive analysis of phonological, morphemic and visual impairments *Cogn. Neuropsychol.*, 1, 1-39.

Seymour, P. H. K., and Porpordas, C. D. (1980). Lexical and non-lexical processing of spelling in dyslexia. In Frith, U. (Ed.), *Cognitive Process in Spelling.* London: Academic Press.

Smith, P. T. (1984). The development of reading: the acquisition of a cognitive skill. In Fletcher, P., and Garman, M. (Eds.), *Language Acquisition.* (2nd ed.) Cambridge: Cambridge University Press.

Soderbergh, R. (1971). *Reading in Early Childhood.* Stockholm: Almqvist and Wiksell.

Sterling, C. M. (1983). Spelling errors in context. *Br. J. Psychol.*, 74, 353-364.

Theios, J., and Muise, J. G. (1977). The word identification process in reading. In Castellan, W. J., Pisoni, D. B., and Potts, G. R. (Eds.), *Cognitive Theory. Vol. 2.* Cambridge, Mass: Lawrence Erlbaum Associates.

Venezky, R. L. (1970). *The Structure of English Orthography.* The Hague: Mouton.

Venezky, R. L., and Johnson, D. (1975). Development of letter-sound patterns in grades one through three. *J. educ. Psychol.*, 64, 109-115.

Venezky, R. L., and Massaro, D. L. (1977). The role of orthographic regularity in word recognition. In Resnick, L. B., and Weaver, P. A. (Eds.), *Theory and Practice of Early Reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Waters, G. S., Seidenberg, M. S., and Bruck, M. (1984). Children's and adult's use of spelling-sound information in three reading tasks. *Memory and Cognition*, 12, 393-395.

Weber, R. M. (1970). A linguistic analysis of first grade reading errors, *Reading Res. Q.*, 5, 427-451.

Wijk, A. (1966). *Rules of Pronunciation for the English Language.* Oxford: Oxford University Press.

Young, D. (1976). *Manual for the SPAR Spelling and Reading Tests.* London: Hodder and Stoughton.

(*Manuscript received 2nd March, 1987*)

## APPENDIX 1

*Friendly words*
same, hand, word, full, bill, slow, mine, seat, rare, barn, bold, sung, boot, cave, lash, suck, hook, cart, mole, brag, sound, space, tried, lower, eight, share, watch, grown, grade, sheep, pitch, clock, trick, spine, bully, greed, fated, stank, miner.

*Unfriendly words*
know, once, free, else, club, neck, jury, busy, poem, pond, ugly, diet, gown, axle, itch, ajar, gulp, rasp, gnaw, snob, major, below, month, wrong, piece, scene, coast, motor, cross, slept, stiff, flood, waist, yacht, dwarf, farce, thorn, clown, daisy.

*Friendly non-words*
bave, brab, fand, filt, flad, gark, hage, mide, mook, nuck, rame, sare, solt, tane, tink, tord, wace, wull, yole, zole, bater, blane, brack, chack, chone, crade, cruss, flack, gight, glave, grafe, jated, mough, prain, rared, shart, slare, stabe, stine, strat.

*Unfriendly non-words*
baft, blid, buch, buke, crem, dirn, frip, grun, jild, plun, saln, shub, slet, spog, swod, tras, trin, vond, wege, woln, aspet, blice, clune, crich, darce, flort, glime, lepen, prent, ronch, rotan, shold, skarn, sotch, stult, sturd, tharn, troct, waral, wrolk.

*Br. J. educ. Psychol.*, 58, 120-126, 1988

# CHILDREN'S LEARNING SKILLS AT THE INFANT AND JUNIOR STAGES: A FOLLOW-ON STUDY

By LEN GREEN AND JEAN FRANCIS
*(School of Education and Humanities, North East London Polytechnic)*

SUMMARY. This follow-on study investigates the learning skills and scholastic attainments of about half of an original cohort of over 2,000 children. The earlier study, when the children were 5 and 6 years old, found observed learning skills to be as predictive of attainment one year later as were IQ scores (Stott *et al.*, 1983a). The follow-up when the children were 9 and 10 years old adopts the earlier experimental design, comparing learning skills with attainments. It also relates early learning skills with attainments four years later. The results and teachers' comments indicate that learning skills strongly influence a child's attainments and that the GCLS (Guide to the Child's Learning Skills) has a contribution to make in the on-going monitoring of children's learning behaviours.

## INTRODUCTION

Since the publication of the Bullock report (DES, 1975) there has been a growing interest in the production of tests or other instruments which are intended to help in the early identification of children with learning difficulties. Some Local Education Authorities have instituted screening procedures to help in the early identification of children with learning difficulties, the intention being to prevent failure before it has actually occurred. The instruments, often locally developed, have not always been found reliable and valid when they have been subjected to rigorous evaluation procedures.

This study focuses on an observational approach to children's behaviours in the learning situation as advocated by Stott (1978). It examines whether, and to what extent, this approach helps teachers to identify junior age children who are likely to have learning difficulties. As a sequel to an earlier study (Stott *et al.*, 1983a) it presents information about the persistence of children's approaches to learning tasks over a four-year interval.

As a result of his work with children of notably poor attainment in reading, Stott (1964) noted that persistent failure could be attributed to various attitudes that interfered with the process of learning. Later, in the Centre for Educational Disabilities at the University of Guelph, he set out to explore poor learning styles as a cause of poor attainment by asking a number of teachers to observe the general styles of coping of those of their 5-year-old children who were likely to run into learning problems. In the school situation the term *learning skills* was found more appropriate than that of *learning styles* in that it moved away from mere description and categorisation and towards the possibility of change. From the teachers' detailed descriptions of the behaviour and attitudes of the children under study, Stott was able to identify 14 categories of faulty or inappropriate learning behaviour. Since these were too lengthy for the assessment of whole classes of children (and, moreover, described only poor learning skills) a shorter checklist was compiled which consisted of seven learning skills likely to be associated with good attainments. This shorter checklist of appropriate learning behaviours became Part One while the 14 inappropriate learning behaviours became Part Two of the GCLS, *Guide to the Child's Learning Skills* (Stott *et al.*, 1983b). Only Part One has been used in the present study.

The first major evaluative study of the GCLS was based on over 2,000 5- and 6-year-olds and reported by Stott *et al.* (1983a). The authors concluded that learning style, as measured by the shorter checklist, could be as equally predictive of future attainment one year later as an IQ score.

There were two hypotheses to be tested in the present study. Firstly, that the manner in which 9- and 10-year-old children cope with learning and problem solving is related to their contemporary scholastic success. Secondly, that learning and problem solving strategies in the early years of schooling are related to future scholastic success. It was possible to test the latter hypothesis in that the children were part of a larger cohort on whom a study had been conducted four years earlier.

TABLE 1

GUIDE TO THE CHILD'S LEARNING SKILLS

Part One: Identification of Learning Skills

Mark the appropriate space against each of the following statements, taking into consideration the child's age.

| | *Certainly applies* | *Applies sometimes* | *Doesn't apply* |
|---|---|---|---|
| Answers in a way that shows attention | | | |
| Settles down well at an activity that needs some concentration | | | |
| Copes with something new without getting nervous or upset | | | |
| Is willing to fall in with the general activities of the class | | | |
| Is willing to try on his/her own | | | |
| Accepts help when a task proves too difficult | | | |
| Is an alert child who enters into activities with interest | | | |

*Scoring should be as follows*

*Certainly applies 0* *Applies sometimes 1* *Doesn't apply 2*

A low score indicates good learning skills

From the *Guide to the Child's Learning Skills* (Stott, Green and Francis, 1983b)

## METHOD

The results presented here relate to those children for whom Part One of the GCLS was completed. Previously, the infant stage of the study had been conducted in 62 infants' schools or departments serving a wide socio-economic range in SW Essex and NE London. At that time the checklists were completed by 221 class teachers in autumn, 1977, for 2,272 children with birth dates between 1st September and 31st December (1,034 5-year-olds and 1,238 6-year-olds). One year later, without feedback of the GCLS results, the children were rated on five-point scales for reading, number and spoken language.

The present study followed on four years later, being conducted with a similar design in 41 of the junior schools receiving the same children. In autumn, 1981, 146 class teachers completed the GCLS Part One for 1,157 children (515 9-year-olds and 642 10-year-olds) from the original cohort. The teachers had no knowledge of the assessments at five and six years. At the beginning of the summer term, 1982, the attainments of 1,128 were assessed on a five-point scale in reading, number and spoken language by the children's class teachers of that time. A further check on children's attainments was made by administering NFER tests later in the summer term to a subsample of 140 children. The aims of the study at the junior stage were:

(1a) To ascertain the test-retest reliability of the GCLS Part One for 9- and 10-year-old children;

(1b) To test the contemporary relationship of the GCLS to teachers' assessments and standardised tests of attainments.

The aims of the follow-on study from the infant to junior stages were:

(2a) To test the constancy of learning skills over a four-year period;

(2b) To test the relationship between the GCLS and attainments four years later.

## RESULTS

### *(1a) Test-retest reliability of the GCLS Part One at 9-10 years old*

Three teachers re-assessed 57 children after four weeks: r=0·81; three teachers re-assessed 59 children after two weeks: r=0·90; two teachers re-assessed 40 children after two weeks: r=0·88; (r=Pearson product-moment correlations, P < 0·001). These reliability coefficients are comparable with those for other checklists of observed behaviour in young children.

### *(1b) Relationship of GCLS scores to teachers' assessments and standardised tests of attainments*

The numbers of children at each GCLS score (0-14) related to their reading attainments are shown in Table 2. A. Kendall's Tau coefficient of −0·47 was significant at the 0·001 level. From this table it is possible to calculate that 74·0 per cent of children with good approaches to learning (GCLS scores 0-1) are found to be above average for reading, with only 4·6 per cent below average. In contrast, only 4·7 per cent of children with the poorest approaches to learning (GCLS scores 8-14) are found to be above average for reading, with 68·6 per cent below average. The distributions for attainments in number and spoken language showed a very similar pattern as that for reading, the coefficients being respectively −0·50 and −0·47.

TABLE 2

GCLS Part One Scores Related to Children's Attainments in Reading

| Numbers of children scoring 0-14 on the GCLS | Assessments for reading | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Very good | | Above average | | Average | | Below average | | Very poor | |
| | N | (%N) | N | (%N) | N | (%N) | N | (%N) | N | (%N) |
| good 0(N= 308) | 130 | 42 | 114 | 37 | 57 | 19 | 7 | 2 | 0 | 0 |
| 1(N= 127) | 29 | 23 | 49 | 39 | 36 | 28 | 13 | 10 | 0 | 0 |
| 2(N= 135) | 22 | 16 | 44 | 33 | 52 | 39 | 15 | 11 | 2 | 1 |
| 3(N= 129) | 17 | 13 | 43 | 33 | 48 | 37 | 19 | 15 | 2 | 2 |
| 4(N= 102) | 5 | 5 | 30 | 29 | 45 | 44 | 20 | 20 | 2 | 2 |
| 5(N= 98) | 3 | 3 | 20 | 20 | 44 | 45 | 28 | 29 | 3 | 3 |
| 6(N= 80) | 2 | 3 | 11 | 14 | 41 | 51 | 24 | 30 | 2 | 3 |
| 7(N= 63) | 2 | 3 | 10 | 16 | 22 | 35 | 19 | 30 | 10 | 16 |
| 8(N= 26) | 0 | 0 | 1 | 4 | 13 | 50 | 9 | 35 | 3 | 12 |
| 9(N= 21) | 0 | 0 | 2 | 10 | 4 | 19 | 8 | 38 | 7 | 33 |
| 10(N= 16) | 0 | 0 | 0 | 0 | 2 | 12 | 7 | 44 | 7 | 44 |
| 11(N= 12) | 0 | 0 | 1 | 8 | 1 | 8 | 6 | 50 | 4 | 33 |
| 12(N= 6) | 0 | 0 | 0 | 0 | 1 | 17 | 3 | 50 | 2 | 33 |
| 13(N= 3) | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 67 | 1 | 33 |
| poor 14(N= 2) | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 100 | 0 | 0 |
| Total (N=1128) | 210 | 18·6 | 325 | 28·8 | 366 | 32·4 | 182 | 16·1 | 45 | 4·0 |

Kendall's Tau= −0·47 (P < 0·001)

Results of tests given to a subsample of 140 children, comparing their GCLS scores to NFER tests and teachers' assessments of attainments, are shown in Table 3. Correlation coefficients are very similar, lying between 0·54 and 0·77.

### *(2a) The constancy of learning skills*

Some constancy between children's scores on the GCLS over a four-year period is shown in Table 4. The correlation of 5 to 9 years is the same as that of 6 to 10 years, but this masks the sex difference. The correlation is higher for girls, r=0·47, than for boys, r=0·33. Table 4 also suggests that there is a slight trend for the mean GCLS scores to attenuate with age, i.e., there is some improvement in learning skills (5-year-old mean GCLS score = 3·45, 9-year-old mean GCLS score = 3·14; 6-year-old mean GCLS score = 3·42, 10-year-old mean GCLS score = 3·01).

TABLE 3

CHILDREN'S GCLS PART ONE SCORES RELATED TO NFER STANDARDISED TEST RESULTS AND TEACHERS' ASSESSMENTS OF ATTAINMENTS (BASED ON SUBSAMPLE STUDIES)

| | | N | r |
|---|---|---|---|
| GCLS Part One scores related to NFER tests | | N | r |
| | Verbal Ability | 140 | −0·59 |
| | Reading | 100 | −0·56 |
| (A) | Maths | 137 | −0·63 |
| | English | 121 | −0·57 |
| GCLS Part One scores related to teachers' assessments of attainments | | | K's Tau |
| | Reading | 140 | −0·62 |
| (B) | Number | 140 | −0·58 |
| | Spoken language | 140 | −0·58 |
| Teachers' assessments of attainments related to NFER attainment tests | | | r |
| | Reading | 100 | 0·59 |
| (C) | Number/Maths | 137 | 0·58 |
| | Spoken language/English | 121 | 0·54 |
| NFER Verbal Ability Test related to NFER attainment tests | | | r |
| | Reading | 100 | 0·77 |
| (D) | Maths | 137 | 0·74 |
| | English | 121 | 0·75 |

*N.B.* r = Pearson product moment correlation.
K's Tau = Kendall's Tau coefficient.
For all coefficients P < 0·001.)
(The numbers vary in that not all of the children were given each test).

TABLE 4

CHILDREN'S GCLS SCORES AS INFANTS RELATED TO THEIR SCORES AS JUNIORS FOUR YEARS LATER

| Infants | Mean GCLS score | Juniors | Mean GCLS score | r |
|---|---|---|---|---|
| 5-yr-olds (N = 515) | 3·45 | 9-yr-olds (N = 515) | 3·14 | 0·40 |
| 6-yr-olds (N = 642) | 3·42 | 10-yr-olds (N = 642) | 3·01 | 0·40 |
| 5&6-yr. boys (N = 578) | 3·79 | 9&10-yr. boys (N = 578) | 3·38 | 0·33 |
| 5&6-yr. girls (N = 579) | 3·08 | 9&10-yr. girls (N = 579) | 2·76 | 0·47 |
| All Infants (N = 1157) | 3·44 | All Juniors (N = 1157) | 3·07 | 0·40 |

(r = Pearson Product Moment Coefficients; P < 0·001 for all correlations)

### *(2b) Relationship between GCLS scores and attainments four years later*

The relationship between children's GCLS scores in the infants' school and their reading attainments four years later is shown in Table 5 (Kendall's Tau = −0·34). The coefficient for number was −0·37, and for spoken language −0·32. For the subsample of 140 children their GCLS scores related to their teachers' assessments of attainments were similar to the foregoing (reading, −0·33; number, −0·40; spoken language, −0·41). Children's GCLS scores were more strongly correlated with NFER standardised test scores (Verbal Ability, −0·45; reading, −0·48; mathematics, −0·49; English, −0·43).

TABLE 5

Infants' GCLS Part One Scores Related to Their Attainments in Reading as Juniors

| Numbers of children scoring 0-14 on the GCLS | Assessments for reading | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Very good | | Above average | | Average | | Below average | | Very poor | |
| | N | (%N) | N | (%N) | N | (%N) | N | (%N) | N | (%N) |
| good 0(N= 234) | 87 | 37 | 81 | 35 | 53 | 23 | 11 | 5 | 2 | 1 |
| 1(N= 159) | 40 | 25 | 61 | 38 | 39 | 25 | 16 | 10 | 3 | 2 |
| 2(N= 143) | 25 | 17 | 46 | 32 | 55 | 39 | 17 | 12 | 0 | 0 |
| 3(N= 133) | 20 | 15 | 41 | 31 | 49 | 37 | 22 | 16 | 1 | 1 |
| 4(N= 91) | 8 | 9 | 28 | 31 | 31 | 34 | 20 | 22 | 4 | 4 |
| 5(N= 96) | 7 | 7 | 30 | 31 | 36 | 38 | 21 | 22 | 2 | 2 |
| 6(N= 75) | 8 | 11 | 12 | 16 | 28 | 38 | 21 | 27 | 6 | 8 |
| 7(N= 49) | 7 | 14 | 3 | 6 | 21 | 43 | 14 | 29 | 4 | 8 |
| 8(N= 54) | 4 | 7 | 7 | 13 | 24 | 44 | 11 | 20 | 8 | 15 |
| 9(N= 38) | 4 | 11 | 10 | 26 | 11 | 29 | 8 | 21 | 5 | 13 |
| 10(N= 27) | 0 | 0 | 5 | 19 | 10 | 37 | 6 | 22 | 6 | 22 |
| 11(N= 15) | 0 | 0 | 1 | 7 | 5 | 33 | 8 | 53 | 1 | 7 |
| 12(N= 9) | 0 | 0 | 0 | 0 | 3 | 33 | 3 | 33 | 3 | 33 |
| 13(N= 3) | 0 | 0 | 0 | 0 | 1 | 33 | 2 | 67 | 0 | 0 |
| poor 14(N= 2) | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 100 | 0 | 0 |
| Total (N=1128) | 210 | 18·6 | 325 | 28·8 | 366 | 32·4 | 182 | 16·1 | 45 | 4·0 |

Kendall's Tau = $-0{\cdot}34$ ($P < 0{\cdot}001$)

From Table 5 it can be calculated that 68·4 per cent of children with good approaches to learning as infants (GCLS scores 0-1) are assessed as above average for reading at the junior stage, while 8·1 per cent are assessed as below average. In contrast, only 20·9 per cent of children with poor approaches to learning (GCLS scores 8-14) are assessed as above average, while 42·6 per cent are assessed as below average.

## DISCUSSION

The investigation involved the completion of the GCLS by teachers at the infant stage and again four years later by a new set of teachers at the junior stage. Such a follow-on strategy does not appear to have been the norm for reliability and validity studies of teacher checklists or rating scales.

The reliability coefficients of the GCLS (0·81 to 0·90) for 9- and 10-year-old children were similar to those obtained with 5- and 6-year-olds and comparable to those of the behavioural items of the Infant Rating Scale (IRS) (Lindsay, 1979). McDermott and Beitman (1984) reported reliability of 0·79 for the Study of Children's Learning Skills, a checklist comprising items from the GCLS. The less than perfect correlation coefficients for reliability studies may relate to the fact that even over a short period of time the learning behaviours of some children may not be stable or teachers' assessments may not be completely reliable.

The relationships between junior children's GCLS scores and their attainments six months later (−0·47 to −0·50) are comparable with those studies which have related children's IQ scores to their attainments. For example, Yule *et al.* (1982) found correlations with the revised WISC of 0·48 for reading accuracy, 0·55 for spelling and 0·70 for mathematics, while Hartlage and Steele (1977) reported correlations of under 0·50 for children's WISC scores with their reading and arithmetic attainments. The GCLS correlated less well with NFER standardised tests of attainment (−0·59) than the NFER Verbal Ability test with the same NFER standardised tests (0·75). However, the NFER tests were given concurrently whereas the GCLS was completed by teachers nine months earlier; during this period changes in family circumstances, health and differential adjustments to school may have produced real changes in the capabilities and learning skills of some children.

Over the infant-junior period the constancy of learning skills was higher for girls (r = 0·47) than for boys (r = 0·33). The overall coefficient (r = 0·40) is similar to that of 0·41 reported by Fogelman (1983) between the behaviour ratings of children in the NCDS cohort at 7 and 11 years. Fogelman asks:

> "To what extent these behaviour ratings reflect meaningfully on the children and how much they reflect real changes in behaviour rather than the efficiency of the rating scales cannot be answered here. What is needed is comparison of these ratings gathered independently by other means. What can be said with assurance from these results is that, in interpreting research studies, they give considerable warning against assuming a static and pathological state for the children identified as deviant."

We agree with Fogelman's note of caution; our concern is that the GCLS Part One should be used by teachers to identify children's current approaches to learning tasks and not as a predictive instrument.

There are few reported studies with which to compare the relationships between children's learning behaviours and their attainments four years later. The GCLS correlations with teachers' assessments (reading −0·34, number −0·37, spoken language −0·32) are less than those for the GCLS with NFER standardised tests (NFER Verbal Ability −0·45, reading −0·48, mathematics −0·49, English −0·49). These coefficients are comparable to those of 0·45 reported by Lindsay (1979) for the IRS (total scores) with the Young Reading Test two and a half years later and other similar studies (Fesbach *et al.*, 1975; Ames, 1977).

A study by Hindley and Owens (1978) which involved the re-assessment of 84 children over a period from 6 months to 17 years demonstrated the variability of IQ scores at different ages. Between 5 and 11 years, 25 per cent changed by over 16 points and a similar percentage between 11 and 17 years changed by 19 points.

From the above it seems that the GCLS as a measure of children's learning skills is about as good a predictor of school attainment as is IQ in young children. However, we believe that the concept of *learning skills* should be given priority as an explanation of children's attainments in school; it is directly induced from manifest behaviour, whereas *intelligence* is essentially unobservable. Keogh (1976) takes a similar view.

> ". . . psychometric, normative evaluations of children may yield less useful educational information than do techniques which get at individual differences in learning strategies, or styles, classroom behaviour, and problem solving techniques."

The coping strategies which children apply to learning situations commend themselves logically as the immediate determinants of success or failure in learning. Behaviour which inhibits or diminishes attention to problems or which interferes with mental processing will understandably hinder learning. The concept of learning skills has the advantage that, being nothing more than a way of describing behaviours, we are not tempted to think of it as a final explanation. On the contrary, we are led to ask in turn what are the determinants of differences in learning skills. In doing so, we have to examine individual motivation, health and neural/sensory intactness and possibly other factors related to the child's environment.

Correspondence and requests for reprints should be addressed to Dr. L. F. Green, 8 Ashburton Avenue, Ilford, Essex, IG3 9ET.

## REFERENCES

AMES, S. G., BECKER, L. D., and DALTON, S. (1977). The predictive validity of the Stanford Early Achievement Test. *Educ. psych. Measur.*, 37, 507-517.

DEPARTMENT OF EDUCATION AND SCIENCE. (1975). *A Language for Life* (The Bullock Report). London: HMSO.

FESBACH, S., ADELMAN, H., and FULLER, W. (1975). The prediction of the reading and related academic problems. Paper presented to the Society for Research in Child Development, Denver.

FOGELMAN, K. (Ed.), (1983). *Growing Up in Great Britain.* Papers from the National Child Development Study. Macmillan for the National Children's Bureau.

HARTLAGE, L. C., and STEELE, C. T. (1977). WISC and WISC-R correlates of academic achievement. *Psychology in the Schools,* 14, 15-18.

HINDLEY, C. B., and OWENS, C. F. (1978). The extent of individual changes in IQ between 6 months and 17 years in a British longitudinal sample. *J. Child Psychol Psychiat.*, 19, 329-350.

KEOGH, B. (1976). Early identification of children at risk for educational programmes. In WEDELL, K., and RAYBOULD, E. C. (Eds.), The early identification of educationally "at risk" children. *Educ. Review*, Occasional Publications No. 6, School of Education, University of Birmingham.

LINDSAY, G. A. (1979). The early identification of learning and the monitoring of children's progress. Unpublished Ph.D. thesis, University of Birmingham.

MCDERMOTT, P. A., and BEITMAN, B. (1984). Standardization of a scale for the study of children's learning styles: structure, stability, and criterion validity. *Psychology in the Schools,* 21, 5-14.

STOTT, D. H., GREEN, L. F., and FRANCIS, J. M. (1983a). Learning style and school attainment *Hum. Learning*, 2, 61-75.

STOTT, D. H., GREEN, L. F., and FRANCIS, J. M. (1983b). *Guide to the Child's Learning Skills* NARE Publications.

STOTT, D. H. (1978). *Helping Children with Learning Difficulties.* London: Ward Lock Educational.

STOTT, D. H. (1964). *Roads to Literacy.* Edinburgh: Holmes McDougall.

WEDELL, K., and RAYBOULD, E. C. (Eds.) (1976). The early identification of educationally "at risk" children. *Educ. Review,* Occasional Publications, No. 6, School of Education, University of Birmingham.

YULE, W., LANSDOWN, R., and URBANOWICZ, M. (1982). Predicting educational attainment from WISC-R in a primary school sample. *Br. J. clin. Psychol.*, 21, 43-46.

(*Manuscript received 20th November,* 1986)

*Br. J. educ. Psychol.*, 58, 127-133, 1988

# SEQUENCE IN HEARING IMPAIRED CHILDREN'S DEVELOPMENT OF CONCEPTS

BY CAROLINE LISTER, CHRIS LEACH AND KAREN WESENCRAFT
*(University of Newcastle upon Tyne)*

SUMMARY. The question of sequence in hearing impaired and hearing children's development of concepts was investigated and a similar order of acquisition of conservation was found regardless of delay. The 24 partially hearing unit, 28 deaf school and 30 hearing ordinary school children, with individual exceptions, yielded the ordering: conservation of number, substance, length, weight, area, volume. Scalogram analyses supported the order with coefficients of reproducibility of $C^1 = 0{\cdot}93$ ($z = 4{\cdot}2$ $P < 0{\cdot}0002$) for the hearing impaired and $C^1 = 0{\cdot}96$ ($z = 4{\cdot}9$ $P < 0{\cdot}0002$) for the hearing. Whilst scalogram analyses clarified the preferred ordering, they also indicated that more than one sequence was significant, and that there was no single fixed order of development.

## INTRODUCTION

The present investigation is concerned with the question of sequence in development of specific concepts. It asks whether hearing impaired children show a similar sequence in their acquisition of conservation concepts to that found in hearing children. Order-of-difficulty evidence is examined. No assumptions are made about underlying structures or stages. The study focuses directly on children's performance when presented with a range of conservation problems involving different attributes. Is there a common order of development of understanding? Does the similar-sequence hypothesis hold in the case of hearing impaired children?

Flavell (1982) wrote that the study of cognitive-developmental sequences has long been recognised as a highly interesting and worthwhile scientific enterprise. He pointed out that even if cognitive development does not prove to be as stage-like as once thought, sequences remain just as important. In fact, he suggested that developmental structure is largely of a sequential vertical rather than stage-like horizontal type and that of primary interest is what develops after what and why. Flavell's comprehensive analyses of structures, stages and sequences in cognitive development and his clear reasons why sequences should figure importantly in theories of development provide background for the present investigation. Also of direct relevance is his examination of the methodological problem of differential test sensitivities and the difficulty in establishing empirically that two or more cognitive acquisitions emerge in a particular order.

Careful empirical investigation is needed to clarify possible sequential regularity. Such investigation assumes that no logically necessary sequence has previously been defined; as Brown and Desforges (1979) argued, "Any hierarchical notion must be examined for conceptual links *prior* to empirical investigation" (p. 88). Brown and Desforges have made clear the complex issues surrounding the question of sequential regularity and no further discussion is needed here. For the present purposes Flavell's (1972) strategy for interpreting cognitive-developmental sequences provides a useful summary of the basic considerations or procedure. Initially, explanation in terms of item-structure logical considerations is tried. If the sequence is not explainable by this alone, a plausible environment-structure explanation is sought. Any explanation in terms of environment-as-experienced which is put forth to account for the order is taken as tentative. If none such can be found, and neither of these two explanations is available or acceptable, *very* tentatively, faute de mieux as Flavell emphasised, there is assumed to be a biological-maturational or organism-structure explanation. Finally, it is recognised that it is not always possible to distinguish between one type of explanation and another and that more than one type may be applicable to a given sequence.

In investigating sequence in development of concepts in hearing impaired and hearing children, it is appreciated that acquiring a concept is not an all-or-none process and that gradations of understanding are evident in relation to conservation. It is also recognised that development of understanding may not be a linear unidirectional phenomenon, that there may be sidewards and backwards movements. Further, there may be many routes to similar understandings, individual differences in sequences of development. Having said all of this, in the well-studied case of conservation some common orderings have been found and the findings have been replicated. Conservation of substance before weight before volume has regularly emerged in research. The place of other attributes, number, length, area, in sequences reported has, however, been less clear. The present research was designed to shed light on this question too. Because of the importance of conservation concepts in the literature and because experiments on understanding of invariance have been so critically examined (see for example Bryant, 1982), concepts of conservation were seen to be appropriate for study in the present investigation.

It has been argued by some Piagetians (e.g., Inhelder *et al.,* 1974) that acquisition of concepts of conservation in a constant order, and similarity in sequence of development, should apply to all children, irrespective of cultural, intellectual or neurological characteristics. Zigler and Hodapp (1986) noted that this universalist argument rests on the belief that such sequences are logically ordered, that the environment promotes development of such sequences, and that the human nervous system may be pre-programmed to develop in this specific way. (But see Brown and Desforges, 1979, and Flavell, 1972, as cited above.) Zigler and Hodapp's concern was specifically with whether a similar-sequence hypothesis should be applicable to *both* familial *and* organically retarded children. Zigler and Balla's (1982) review of research into the developmental-difference controversy had included a comprehensive examination of order-of-difficulty evidence and studies of conservation. They had concluded that despite the diversity of methodologies and sample characteristics, the data showed rather consistent support for the similar-sequence hypothesis. Although the above research has been concerned with mentally retarded children in particular, it raises issues of fundamental importance in relation to the development of children with other handicapping conditions.

In the present investigation, the similar-sequence hypothesis of the developmental theorists (e.g., Zigler and Balla, 1982; Zigler and Hodapp, 1986) was extended to work with the hearing impaired. Such an extension, as far as we know, has not previously been investigated.

A question which *has* been investigated but which is as yet unresolved is that of delay in hearing impaired children's development of understanding of conservation. Markedly different findings have been reported. In some studies delays in understanding of conservation have been found to be as great as six or seven years; in others no significant delay has been found. (See, for example, Oleron and Herren, 1961; Furth, 1964, 1966, 1973; Darbyshire and Reeves, 1969; Watts, 1976, 1981, 1982; Rittenhouse and Spiro, 1979.) The differences between the investigations, in presentation, preliminary teaching, tasks used, children's ages and other variables, make it difficult to compare studies and draw conclusions (See Lister *et al.,* 1987). Not only are differing procedures, materials and subjects likely to affect results but also criteria used in interpreting responses may differ and give rise to conflicting conclusions about when understanding of conservation develops. The problem of criteria, what should be required or what is appropriate to expect in work with hearing impaired children, was given particular attention in the present investigation.

In summary, sequence in hearing impaired children's development of concepts was the main question studied. Testing the similar-sequence hypothesis involved presenting both hearing impaired and hearing children with the same tasks and analysing their responses using the same criteria to ensure comparability. Such a procedure can also shed light on the question of delay.

## METHOD

### *Subjects*

Twenty-four partially hearing unit children, 28 deaf school children, and 30 hearing ordinary school children were seen individually in the familiar surroundings of their own schools. The partially hearing unit children (13 boys and 11 girls) were aged from seven years nine months to 16 years four months (mean 12 years eight months) and had hearing losses (average loss in the better ear) from 40dB to 105dB (mean 77dB). The deaf school children (24 boys and four girls) were aged four years six months to 16 years four months (mean 11 years eight months) and had hearing losses from 69dB to 119dB (mean 97dB). The hearing ordinary school children (17 boys and 13 girls) were aged four years nine months to 11 years two months (mean eight years 10 months). The schools were all located in the same urban area.

It was expected that the age range of the hearing sample, four to 11 years, would cover total non-conservers to total conservers and that a hearing impaired sample aged four to 16 years would do the same. The hearing impaired sample included all available children other than those with additional handicaps.

*Procedure*

The conservation tasks used were, with some modifications, like those used by Piaget (See Piaget and Inhelder, 1941; Piaget, 1952; Piaget *et al.,* 1960). In essence, each problem began by presenting two sets of material, A and B, which the child was able to judge equal with regard to the particular attribute in question; then one set, B, was transformed in full view to give $B^1$ and the child asked to compare A and $B^1$ with regard to the same attribute. Cases of inequality were included for each attribute in order to clarify the child's understanding and expose any tendency to agree to sameness regardless of the actual situation. After a child made a judgment, the experimenter asked the child for an explanation. The child's responses were tape-recorded as well as written down and subsequently analysed in detail.

The specific problems were, briefly, as follows. *Number:* eight and then 12 beads were transferred into different-shaped glass containers; then 10 pencils were taken out of their jars and bunched nearer together and spread farther apart than the jars. *Substance:* glass containers, different from those used above, were used for liquid transformations; then plasticine transformations were made. *Length:* sticks were "staggered" as well as placed between Muller-Lyer arrowheads; then strips of card were cut and positioned in different ways. *Area:* a small number of houses on each of two fields was presented and one arrangement transformed and then the same was done with larger numbers of houses; then, Piaget's rearranged shapes were incorporated into the problem of potato plots on fields and questions of complementary as well as interior area were asked. *Weight:* plasticine shapes were transformed; then plastic bags of rice or sand were transformed. *Volume:* Piaget's interior and displacement volume problems were presented; then a displacement case was introduced in which there was an actual transformation of material — wooden and metal unit-cubes were used in these problems.

The tasks for conservation of the different attributes were presented in a constant order, as above. The order was not randomised because the effect of facing any hearing impaired child with a volume conservation task, or other particularly difficult task, *first*, was thought likely to sap confidence and reduce interest.

Particular care was taken to ensure clear communication, understanding of the questions being asked and the attributes being considered. Both oral and manual communication were used as appropriate with the hearing impaired children.

To qualify as conserving an attribute a child had to recognise conservation in *every one* of the problems concerned with that attribute, making consistent correct judgments. Explanations for judgments were *not* required, to be classed as a conserver, for the following reasons. Researchers have disagreed about what criteria should be employed in assessing responses to conservation problems. Accepting judgment alone may result in "false positives" but requiring explanation in addition to correct judgment may result in "false negatives". Because there has been disagreement about criteria even in work with ordinary hearing children and because in work with hearing impaired children the risk of "false negatives" if explanation is required is still greater, unreasonably great, it was decided that the appropriate criterion for conservation was consistent correct judgment.

Thus, for each attribute each child was classified as either conserving (correct for *every* problem) or not conserving (incorrect for one or more problems). Whilst intermediate classification was considered in order to clarify certain findings or interpretations, for the analyses each child was classified as either conserving or not conserving.

## RESULTS

When the above criteria were used to classify each child as conserving (+) or not conserving (−) on each attribute, the overall outcome was as shown in Table 1. Seventeen distinct patterns in the hearing impaired children's recognition of conservation are displayed, and eight patterns in the hearing children's recognition of conservation. Table 1 gives clear evidence for a developmental sequence.

In Table 1, the partially hearing unit and deaf school samples have been combined to form the hearing impaired sample. There was overlap in the two samples' degree of hearing loss and there were no observed characteristics of the samples which would indicate that it would be inappropriate to combine them in analysing the data.

TABLE 1

SUMMARY OF RESULTS SHOWING DEVELOPMENTAL SEQUENCE RELATED TO ATTRIBUTE: NUMBERS OF HEARING IMPAIRED (N = 52) AND HEARING CHILDREN (N = 30) SHOWING EACH PATTERN

| | Number | Substance | Length | Weight | Area | Volume | Hearing Impaired | Hearing |
|---|---|---|---|---|---|---|---|---|
| Pattern a | – | – | – | – | – | – | 9 | 2 |
| Pattern b | + | – | – | – | – | – | 12 | |
| Pattern c | + | + | – | – | – | – | 3 | 3 |
| Pattern d | + | + | + | – | – | – | 3 | |
| Pattern e | + | + | + | + | – | – | 3 | 3 |
| Pattern f | + | + | + | + | + | – | 1 | 11 |
| Pattern g | + | + | + | + | + | + | 2 | 4 |
| Pattern h | + | + | + | – | + | – | 1 | 3 |
| Pattern i | + | + | – | + | – | – | 1 | 3 |
| Pattern j | + | – | + | + | + | – | 2 | |
| Pattern k | + | – | + | – | – | – | 5 | |
| Pattern l | + | – | – | + | – | – | 3 | |
| Pattern m | – | + | – | – | – | – | 2 | |
| Pattern n | – | + | + | – | – | – | 1 | |
| Pattern o | – | – | – | + | – | – | 2 | |
| Pattern p | – | + | – | – | + | – | 1 | |
| Pattern q | – | + | – | + | – | – | 1 | |
| Pattern r | – | + | + | + | + | – | | 1 |
| Number conserving attribute: | | | | | | | | |
| Hearing impaired | | | | | | | | |
| | 36 | 19 | 18 | 15 | 7 | 2 | 52 | |
| Hearing | | | | | | | | |
| | 27 | 28 | 22 | 22 | 19 | 4 | | 30 |

Scalogram analyses of the data in Table 1 support the order conservation of number substance length weight area volume. For the hearing impaired sample, 33 of 52 children showing patterns a, b, c, d, e, f, and g are consistent with the order, and for the hearing, 23 of 30 children with patterns a, c, e, f, and g comply with the order. Using Goodman's (1959) method of testing whether the order defined in Table 1 is reliable better than chance, Sagi's (1959) coefficient of reproducibility for the hearing impaired sample is $C^{1} = 0{\cdot}93$, a value that is highly significant ($z = 4{\cdot}2$, $P < 0{\cdot}0002$) and for the hearing sample is $C^{1} = 0{\cdot}96$, also highly significant ($z = 4{\cdot}9$, $P < 0{\cdot}0002$). This indicates that the performance of both groups is consistent with the order given.

For the combined hearing impaired sample, the preferred sequence was clearly that shown in Table 1. The same was true for the partially hearing unit sample taken on its own. For the deaf school sample, the Sagi coefficient of reproducibility for the ordering shown in Table 1 was 0·93 while the only higher coefficient for an alternative ordering was 0·95. Since there was no clear distinction to be made between them, the sequence in Table 1 was arguably the best estimate of order of difficulty or sequence of development.

In the hearing impaired sample, there were marked individual differences not only in patterns of conservation but also in ages at which children developed conservation concepts. Figure 1 shows quartile plots for the age-attribute data for both hearing impaired and hearing children. For each attribute, these plots give the median age for those children conserving (the dots in the middle of the plots), the interquartile range (the gaps around the median), and the range (the lines extended to the lowest and highest ages). See Leach (1988) or Seheult (1986) for information on quartile plots. The plots here indicate the extent of delay in hearing impaired children's recognition of conservation and its consistency across attributes. The difference of three to four years between the median age at which the hearing and the hearing impaired samples achieved conservation is consistent across attributes.

The youngest hearing impaired child to conserve all six attributes was 13 years two months and the oldest who did not recognise conservation of any attribute was 14 years nine months. Of the hearing children, the youngest to conserve all six attributes was nine years one month and the oldest who did not conserve any attribute was five years 0 months.

## DISCUSSION

Overall the findings show that hearing impaired children develop understanding of concepts in a similar order to that found in hearing children. More exactly, when the same conservation tasks are presented and the same criteria used in analysing responses, the partially hearing unit, deaf school and ordinary school children reveal similar sequences in development. Whilst there are minor discrepancies and individual exceptions, a common developmental order, highly significant for both hearing impaired and hearing samples, is the following: conservation of number, substance, length, weight, area, and finally volume. For the hearing impaired sample, this order is strikingly clear. For the hearing sample, there is a slight suggestion of a substance-number reversal, but with 28 children who conserve substance and 27 number it is not possible to draw any firm conclusion on this.

The answer to the main question investigated appears clear. Hearing impaired children do develop concepts of conservation of quantity in a similar sequence to that found in hearing children. That is not to say that development of conservation concepts occurs in a single or fixed order. Scalogram analyses clarified preferred orderings but also indicated that more than one sequence may be highly significant. It would seem likely that different orderings for some, if not all, of the attributes could be found.

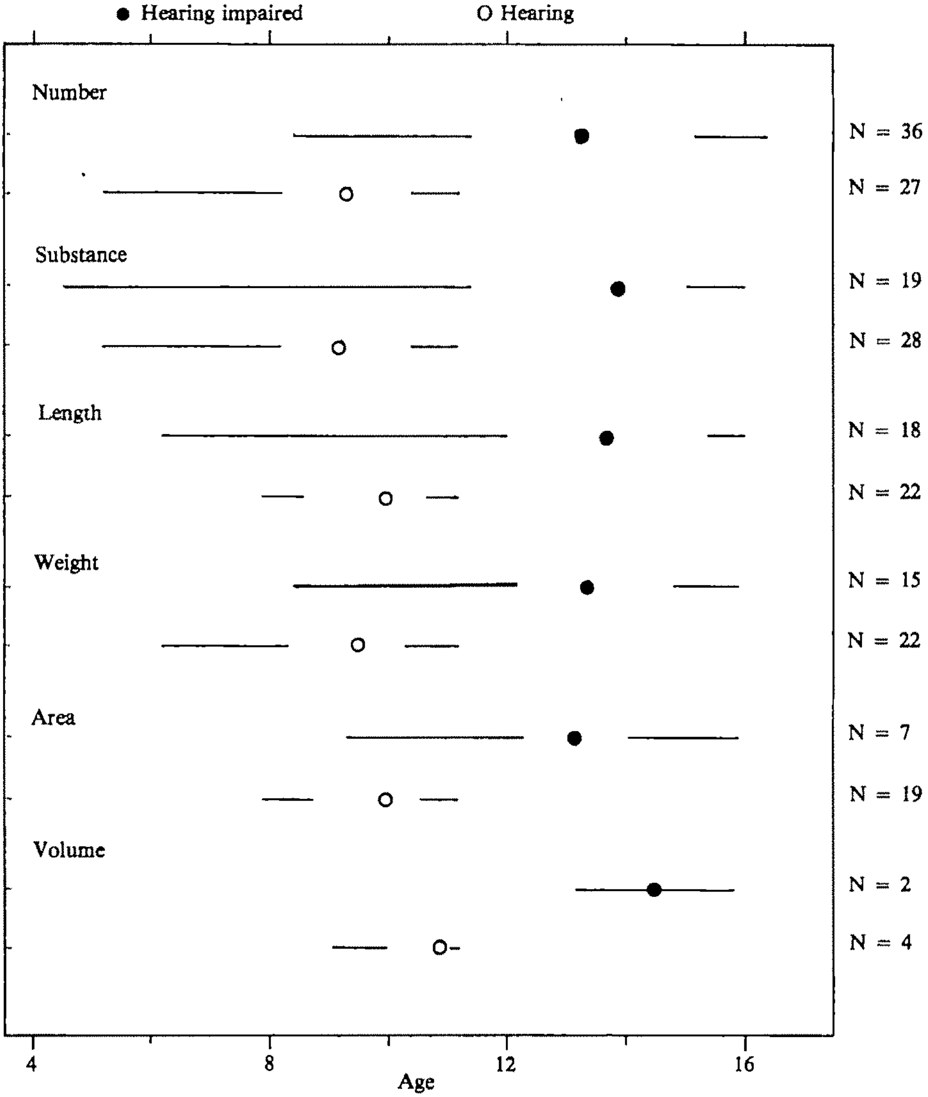
No single order for all six had been defined as logically necessary. Different environments and experiences are likely to influence or alter order of difficulty, as are organism or individual differences. Goldschmid and Bentler (1968) wrote, "The problem of relative difficulties of different tasks of conservation has unfortunately not been clarified by the results of our studies. The rank orders of difficulties of various tasks are not identical from scale to scale, or from sample to sample, so that conclusions regarding relative difficulties of tasks cannot be drawn from our data. There appears to be no one sequence of task difficulties, making it impossible to establish *the* Guttman sequence" (p. 801). The findings in the present investigation are in accord with Goldschmid and Bentler's, and, as Brown and Desforges (1979) pointed out, "In many studies there are numerous children whose performance does not conform to the perfect Guttman scale" (p. 65).

The present investigation also shed light on the question of delay in hearing impaired children's development of understanding of conservation. When the same tasks were presented and the same criteria for conservation used, the hearing impaired children showed a lag of three to four years compared with hearing children. The delay was quite consistent across all six attributes and there was little difference between the partially hearing and deaf groups in degree of delay.

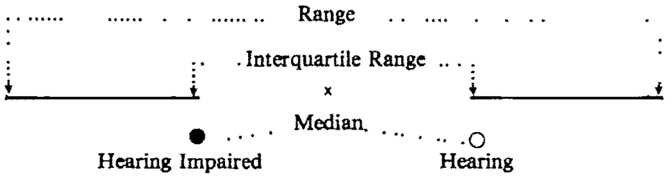
The marked individual differences in the hearing impaired children's patterns of conservation need some comment. They are not easy to explain. One possible explanation may lie in the nature of hearing impaired children's development of understanding. An alternative explanation may be that as chronological age has systematic effects on children's general developmental level, with a wide age range, as in the case of the hearing impaired children in

FIGURE 1

QUARTILE PLOTS OF AGES OF HEARING IMPAIRED (●) AND HEARING (O) CHILDREN CONSERVING EACH ATTRIBUTE



Key to Figure 1



this study, one can expect a wide range of individual differences which in turn will be reflected in the patterns produced by the scalogram analyses. (We are grateful to a referee for suggesting this second possibility.)

To conclude, the extension of the developmental theorists' similar-sequence hypothesis to cover children with other handicapping conditions proved fruitful, revealing that these children, regardless of any delay, may show the same order in their development of concepts as ordinary children.

Correspondence and requests for reprints should be addressed to Dr. Caroline Lister, Educational Psychology Advanced Course, 4th Floor, Department of Psychology Building, University of Newcastle upon Tyne, Claremont Place, Newcastle upon Tyne, NE1 7RU. Karen Wesencraft is at the same address. Chris Leach is at the Department of Psychology, University of Newcastle upon Tyne.

## REFERENCES

Brown, G., and Desforges, C. (1979). *Piaget's Theory: A Psychological Critique.* London: Routledge and Kegan Paul.

Bryant, P. (Ed.) (1982). *Piaget: Issues and Experiments.* Leicester: British Psychological Society.

Darbyshire, J. O., and Reeves, V. S. (1969). The use of adaptations of some of Piaget's tests with groups of children with normal and impaired hearing. *Br. J. Disorders Communic.,* 4, 197-202.

Flavell, J. H. (1972). An analysis of cognitive developmental sequences. *Genet. Psychol. Monogr.,* 86, 279-350.

Flavell, J. H. (1982). Structures, stages, and sequences in cognitive development. In *15th Annual Minnesota Symposium on Child Psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Furth, H. G. (1964). Conservation of weight in deaf and hearing children. *Child Dev.,* 35, 145-150.

Furth, H. G. (1966). *Thinking Without Language.* New York: Free Press.

Furth, H. G. (1973). *Deafness and Learning: A Psychosocial Approach.* Belmont, Calif.: Wadsworth.

Goldschmid, M. L., and Bentler, P. M. (1968). The dimensions and measurement of conservation. *Child Dev.,* 39, 787-802.

Goodman, L. R. (1959). Simple statistical methods for scalogram analysis. *Psychometrika,* 24, 29-43.

Inhelder, B., Sinclair, H., and Bovet, M (1974). *Learning and the Development of Cognition.* London: Routledge and Kegan Paul.

Leach, C. (1988). Guidelines for data presentation. In Sternberg, R. S. (Ed.). *The Psychologist's Companion: A Guide to Scientific Writing for Students and Researchers.* New York: Cambridge University Press; Leicester: British Psychological Society.

Lister, C., Charlton, J., Madden, L., and Leach, C. (1987). The development of understanding of conservation in hearing impaired children. *Early Child Development and Care,* 27, 493-505.

Oleron, P., and Herren, H. (1961). L'acquisition des conservations et le langage: etude comparative sur des enfants sourds et entendants *Enfance,* 14, 203-219.

Piaget, J. (1952). *The Child's Conception of Number.* New York: Humanities.

Piaget, J., and Inhelder, B. (1941). *Le Developpement des Quantites Physiques chez l'Enfant.* Neuchatel: Delachaux et Niestle.

Piaget, J., Inhelder, B., and Szeminska, A. (1960). *The Child's Conception of Geometery.* London: Routledge and Kegan Paul.

Rittenhouse, N., and Spiro, R. J. (1979). Conservation in deaf and normal-hearing children. Paper presented at the annual International Convention for Exceptional Children, April.

Sagi, P. (1959). A statistical test for the significance of the coefficient of reproducibility. *Psychometrika,* 24, 19-27.

Seheult, A. (1986). Simple graphical methods for data analysis. In Lovie, A. D. (Ed.). *New Developments in Statistics for Psychology and the Social Sciences.* Leicester: British Psychological Society; London: Methuen.

Watts, W. J. (1976). Deafness, thought and language: psychological and educational implications of deafness. Unpublished Ph.D. thesis, University of Sussex.

Watts, W. J. (1981). The performance of deaf, partially hearing and normally hearing children on conservation tasks of cardinal number, discontinuous quantity and length. *Teacher of the Deaf,* 5, 11-19.

Watts, W. J. (1982). The performance of deaf, partially hearing and normally hearing children on conservation tasks of weight and area. *Teacher of the Deaf,* 6, 5-9.

Zigler, E., and Balla, D. (Eds.) (1982). *Mental Retardation: The Developmental-Difference Controversy.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Zigler, E., and Hodapp, R. M. (1986) *Understanding Mental Retardation.* Cambridge: University Press.

(*Manuscript received* 8th *May,* 1987)

# BOOK REVIEWS

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., and Tarule, J. M. (1986). *Women's Ways of Knowing*. New York: Basic Books, pp. 256, £19·95 hbk. ISBN 0-465-09212-8.

Just as Carol Gilligan expanded our understanding of moral development by including women's conceptions of responsibility and moral reasoning, the authors of this book claim to have enlarged our understanding of intellectual development by listening to women. Traditionally in both those areas of development descriptions of the male experience have been used to define the human experience. In demonstrating women's intellectual competence the focus has tended to be on what has traditionally been perceived as male characteristics such as the development of abstract, impersonal, formal thought culminating in the ability to apprehend universal laws of nature and justice. By comparison those that dealt with the personal and interpersonal were regarded as emotional and largely relegated to women. Most institutions of higher education whose stated mission is to promote intellectual development have been designed, developed and run by men and continue to use male models of knowing which may be at odds with women's educational needs. The central aim of the research which formed the basis of this insightful and thought-provoking book was to improve our understanding of the needs of women learners in the post-secondary education stage in order to design an education appropriate for them.

One hundred and thirty-five women from all walks of life were interviewed to trace their natural ways of thinking about thinking and about themselves as knowers and to see how such notions evolved over time. The researchers were particularly interested in seeing how teaching approaches might support or undermine women's ways of knowing. They also wanted to know about women's experiences as mothers because (a) teaching the next generation is traditionally regarded as women's central role, and (b) the wisdom women have gained through maternal practice and maternal thinking could be particularly illuminating to educators intent on promoting intellectual development.

Analysis of these data revealed five different ways of knowing or epistemological positions that women use to make sense of their experiences. They are described as silence, received knowledge (the voice of others), subjective knowledge (the inner voice), procedural knowledge (the voice of reason) and constructed knowledge (integrating the voices). The importance of "gaining a voice" (dialogue and interaction) was a recurring theme used by the women to depict their experience of development. The development of the intellect and the capacity for dialogue were perceived to be intertwined.

The women who saw things from the first of these positions seemed unaware of the power of language for transmitting and receiving knowledge and consequently lived in a silent, isolated world controlled by others. Predictably, perhaps, they came from families where one parent — usually, but not always, the father — ruled over the others. The women in the next position saw themselves as capable of receiving ideas but not creating them. Thus they too were subjected to the standards, directions and authority of others. Typically these women's parents had done most of the talking and the children had done most of the listening. Parental authority — even during adolescence — was rarely questioned. The politics of family talk associated with these two positions can be compared with Paulo Friere's "banking model" in which the role of the educator is to fill the learner by depositing information. The learner is not required to think or question that received information but to he heavily dependent upon it.

When women become conscious of their voicelessness they begin to shift their orientation from listening to the voices of others to listening to their own inner voice. Characteristically women in this position distinguished between truth as feeling which comes from within and truth as ideas which comes from without. Women who become aware that their intuitions could be wrong worked hard to cultivate a voice of reason. To them knowing had become associated with careful observation and analysis. They become absorbed in the business of acquiring and applying procedures for obtaining, reflecting on, evaluating and communicating knowledge. They sought to establish the truth by objective dispassionate means. These, the authors argue, are procedures for separate knowing and are based on a male model of knowing. The teaching approach which cultivates them is essentially an adversarial form of reasoning — one which induces cognitive conflict in the students' minds. All the women interviewed found this educational experience debilitating rather than energising suggesting that as an approach it may be particularly inappropriate for women. The evidence suggests that most women prefer and need an education in which dialogue takes precedence over observation, acceptance and understanding over judgment and assessment, and collaboration over debate. Connected knowing is the phrase used by the authors to characterise the kinds of procedures required for constructed knowledge. These are cultivated in the contexts of collaborative work where dialogue is valued rather than through didactic and adversarial approaches to teaching where the impersonal and abstract are required ways of reasoning.

This book succeeds in throwing some light on why even well-intentioned educators and parents leave women reluctant to speak up for what they want and from that point of view should be read by all educators. Whether the findings are uniquely female must, however, remain an open question.

FLORA MACLEOD

## REFERENCES

FREIRE, P. (1971). *Pedagogy of the Oppressed.* New York: Sea View.

GILLIGAN, C. (1982). *In a Different Voice: Psychological Theory and Women's Development.* Cambridge, MA: Harvard University Press.

---

ENTWISTLE, N. (1987). *Understanding Classroom Learning*. London: Hodder and Stoughton, pp. 134, £5·95 pbk. ISBN 0-340-40736-0.

For decades teachers, psychologists and a host of others have been attempting to understand the complexities of classroom learning and Entwistle is no exception. In his book he undertakes the difficult task of summarising learning theories and using them to present an interactive model of classroom processes.

The book — which is aimed at teachers, students and administrators — provides a clear and concise account of a wide variety of theories. It is cheap, easy to read and the sort of book students might want to refer to time and again as they gain in teaching experience. Through the use of examples Entwistle has succeeded in illustrating how complex research findings apply to the classroom context and yet rarely does he appear patronising to his reader. The layout of the book and the model presented in the final chapter allow for easy cross-referencing and a reminder of the evidence considered.

As a student text this book has much to offer. It does not, however, present anything very new or exceptional; nor indeed does it provide a complete review of pertinent research in the field of classroom learning. More specifically, for example, no mention is made of Doyle's work (e.g., 1977, 1981, 1983 and 1986). Entwistle may disagree with Doyle's thinking but since it culminates in what many (e.g., Bennett *et al.*, 1984 and Desforges, 1985) consider to be a convincing model of classroom processes it should not be overlooked in what is described as "... an attempt ... to identify (the) key concepts and theories which are most likely to speak to the classroom experience of teachers" (vii).

To summarise, Entwistle's book is a good introduction to some of the most useful theories of classroom learning. It does not, however, provide a comprehensive review of recent, relevant research, nor does it provide the key to classroom processes for which so many of us are searching.

ANNE D. COCKBURN

## REFERENCES

BENNETT, S. N., DESFORGES, C. W., COCKBURN, A. D., and WILKINSON, B. (1984). *The Quality of Pupil Learning Experiences.* London: Lawrence Erlbaum Associates.

DESFORGES, C. W. (1985). Matching tasks to children. In BENNETT, S. N., and DESFORGES, C. W. (Eds.), *Recent Advances in Classroom Research.* Edinburgh: Scottish Academic Press.

DOYLE, W. (1977). Paradigms for research on teacher effectiveness. In SHULMAN, L. (Ed.), *Review of Research in Education.* Itasca, Ill.: F. E. Peacock.

DOYLE, W. (1981). Research on classroom contexts. *J. Teacher Educ.*, 32, 3-6.

DOYLE, W. (1983). Academic work. *Rev. educ. Res.*, 53, 159-200.

DOYLE, W. (1986). Classroom management and the curriculum: a strategic research site. In WITTROCK, M. C. (Ed.), *Handbook of Research on Teaching* (3rd ed.). New York: Macmillan.

ORTON, A. (1987). *Learning Mathematics: Issues, Theory and Classroom Practice.* London: Cassell, pp. 179, £7·95. ISBN 0-304-31369-6.

This book is intended to provide "an introduction to the study of mathematics that raises issues (and) provides discussion of theory and practice at a relatively elementary level" for "all courses which include studies of learning mathematics". Such ambition recommends it as a valuable contribution to teacher education, yet it falls substantially short of this because the emphasis is on easy-to-read versions of established theoretical work rather than on providing challenging invitations to analyse their applicability to classroom events.

This approach is largely determined by the way in which the book is organised: whereas it is an attractive idea to focus on a series of questions (about, for instance, the cognitive demands of learning mathematics; the problem of sequencing; readiness; discovery) and to attempt an eclectic approach to the discussion of theory, this requires either an assumption that the range of theories are known to the reader or that appropriate detail be provided. Given the intended audience, the latter course is reasonable — but the decision to supply this detail during discussion of the question has the inevitable consequence that selected aspects are delivered in destructively abbreviated form and the reader is consigned to passive acceptance of that particular selection. Unoriginal as it may have seemed to the author, a chapter on each of the major orientations would have been more appropriate, followed by what would then have been a welcome conversion of these ideas into competing frameworks for the classroom examples which are a significant feature of the book. There is evidence that this would have been successful: in the discussion of "readiness", concentrated very largely on Piaget, we are treated to an unusually coherent elementary review of the work and the questions it raises.

The weakness of the book is most apparent in its least successful chapter: "Can pupils discover mathematics for themselves?" The chapter begins with a description of an "assignment" requiring the determination of the number of diagonals of each of the polygons. The author is apparently unaware of the manner in which this is typically presented by competent practitioners because, rather than entering what could have been a fruitful discussion of the insights offered by Vygotsky, Bruner and the "schema" theorists, we are offered little more than the assertion that "the entire assignment may be given to children to work on for themselves".

This same chapter continues to evidence lack of familiarity with contemporary practice in statements which can only be considered as dangerous for the student: "simply playing" with coloured rods "takes the children a very long way to the mastery of number relationships"; "It is questionable, though, whether a child who cannot discover anything can benefit from expository teaching methods"; "a certain vagueness about terms such as 'discovery', 'investigation' and 'problem solving' . . . does not matter all that much".

Further, an unfortunate misunderstanding of the potential of an activity focused on the number of rectangles on a 5 × 3 pegboard is apparent in calling it "free and exploratory" without qualification and without any attempt to place it in the context of the variety of theoretical perspectives. And to encourage the assertion that "free and exploratory investigation" is acceptable if the "results do not constitute essential knowledge" is less than helpful.

The saddest omission is any adequate discussion of the role of language in making sense of mathematics, and of the implications for classroom activity of a deliberate policy of using oral work. It is a poor choice of chapter heading to ask "Does language interfere with mathematics learning?", for although there are some useful observations here it certainly does not provoke a consideration of the role of language.

There is a need for recent developments in mathematics education, including the received truths of the Cockcroft Report and the subsequent attempt to make sense of it, to be subjected to an analysis by teachers who are able to bring theory to bear on practice. It is unfortunate that the confidence in the first chapter that "teachers need theories" is followed by a text which leaves the uneasy feeling that those which are available are irrelevant.

RICHARD DUNNE

Osborn, A. F., and Milbank, J. E. (1987). *The Effects of Early Education: A Report from the Child Health and Education Study*. Oxford: Clarendon Press, pp. 294, £22·50 hbk., ISBN 0-19-827801-2; £7·95 pbk., ISBN 0-19-827800-4.

It is part of the human condition that there should be more difficult questions than simple answers; but here is a book that has at least one straight answer to one important question: can preschool education have beneficial effects on children? "The simple answer is, indubitably, yes" (p. 237).

The full answer is, of course, a little more complicated. Nevertheless, after subjecting their wealth of data to severe Popperian onslaughts, in attempts to destroy associations between preschool attendance and cognitive gains at ages 5 and 10, Osborn and Milbank confidently maintain the robustness of their central finding: "Preschool education provided by ordinary nursery schools and playgroups in Britain can have a positive effect upon the cognitive development of the children who attend them" (p. 238). Preschool education is used here as an umbrella term to refer to seven different kinds of institutions; three maintained (local authority day nurseries, local education authority nursery schools and classes) and four independent (home playgroups, hall playgroups, day nurseries and nursery schools). The evidence shows that even though these institutions are intended to serve different purposes, offer different kinds of experience, and are staffed by adults with different training, there are few statistically significant differences between the test performances of children who attend them. The important differences are between children who receive preschool education and those who do not.

So far, so simple. But in a study based on a sample of 9,000 children (drawn from the Child Health and Education Study cohort of 1970), linked to a study of 17,000 preschool institutions, there is bound to be a little more fine shading. With irreproachable clarity, Osborn and Milbank guide the reader through the procedures of the study and the intricate process of statistical analysis and adjustment. The small print reveals important subsidiary findings. First, there is substantial social inequality in access to preschool education: "As many as 46 per cent of the most disadvantaged children had received no form of preschool education compared with only 10 per cent of the most advantaged group" (p. 206). Secondly, the achievement of children attending nursery schools was enhanced relative to children who went to nursery classes. Thirdly, children who start infant school early, at the age of 4, achieve no higher test scores than children who start at the age of 5. Fourthly, the children who attended home playgroups had the highest mean scores in four out of seven tests, and the second highest in the other three.

It does not escape the authors' notice that these results have far-reaching implications and they go on to explore the social and educational meanings of their work. However, they are not in the business of writing a manifesto for the expansion of preschool education; they have remained true to their trade as dispassionate researchers. As a result, their discussion lacks the urgency of the reformer, and the energy of political commitment: this is perfectly proper behaviour, but it will disappoint many readers.

They are decidedly firm, for example, about the need to review practice and provision in nursery classes and reception classes containing 4-year-olds, and repeatedly underline their view that a nursery curriculum is the most appropriate for under-5s. However, since Osborn and Milbank laid down their pens, patterns of admission have changed dramatically; an estimated 80 per cent of the country's 4-year-olds are, at present, in infant or primary schools, so that it is now even more pressingly important that the policy-makers should be made to sit up and take notice of what this study tells us. But the authors seem to refrain from driving their arguments home, and their lack of enthusiasm may lead to a lack of audience response.

In the same way, Osborn and Milbank's warning that the positive effects of playgroup attendance should not be interpreted as a justification for increased financial support for playgroups, to the detriment of the more expensive day nurseries and nursery schools, may also go unheeded. I fear that those with power over preschool purse-strings will not readily be moved by rational and temperate research-based recommendations. If the authors' scrupulous level-headedness were only fired by a little more fervour, there might be more chance of their work having some effect.

Longitudinal studies have inherent defects, as the authors are willing to admit; there is inevitably a certain loss of topicality, since the issues of today are different from those of 1975 or 1980. So, for example, the exciting growth of joint centres for under-fives is not considered here, and there is a questionable slant to the investigation into the achievements of children from ethnic minority groups. Children of Afro-Caribbean origin and Indian/Pakistani origin are considered in a section headed "Children with special educational needs" and referred to as "the most vulnerable children" (p. 111). However, the results of this investigation are encouragingly positive (and should be publicly broadcast in Dewsbury): pre-school institutions attended by ethnic minority group children provide a better learning environment for *all* the children who attend them than do those with no such children.

The most disappointing section of the book is the brief analysis of curriculum: the authors admit the limitations of a questionnaire approach to an assessment of practice, but the four scales they used to monitor provision in the 17,000 institutions seem unnecessarily flimsy. The "finding out" scale, for example, measures the availability of just four items: sand, water, pets, and an interest/discovery table. These activities hardly constitute an adequate description of what might be provided in such an important area of learning and experience. The conceptual poverty of the instruments used to describe curriculum matters may account for the absence of any worthwhile conclusions to be drawn from them. On the other hand, there is an unexpected bonus in the authors' contribution to the continuing High/Scope debate. In a frisky aside, Osborn and Milbank raise the politest of question-marks over Weikart's claims for the Perry Preschool Project:

> "We suggest that the most ardent advocate of preschool education would not expect attendance at even the very finest preschool nursery to increase the educational potential and life chances of children to the extent suggested . . . there must have been contributory factors additional to the preschool intervention" (p. 201).

Even simple answers raise further questions: and this study implicitly raises a great many, some of which are uncompromisingly political. Who will be prepared to take the study seriously, and implement its recommendations? Who will provide the money to do so? Who will take the lead in facing up to the glaring social inequalities in preschool provision? (A reminder: "The children most in need of preschool education were the ones who were most likely to have had none" p. 92.) Will this finding be taken as official endorsement for the concept of preschool education as rescue for the underprivileged of society? Will the authors be seen as suggesting that there is a single "best buy" for preschool education? What will playgroup workers, nursery assistants and teachers do to reconcile their sometimes exclusive and elitist attitudes towards other preschool workers with the finding that type of preschool attended has little significant effect on children's achievement? Will this study stimulate others to undertake a more critical and penetrating investigation of what actually happens to children inside the preschool institutions they attend? How long will it be before we know the answers to these questions?

MARY JANE DRUMMOND

---

SHARRON, H. (1987). *Changing Children's Minds.* London: Souvenir Press, pp. 336, £16·95. ISBN 0-285-65034-3.

Over the past five years the interest amongst educators, psychologists and, more latterly parents in the ideas of Reuven Feuerstein and his colleagues, and their practical implications for assessment and remedial and special education, has grown enormously. Unfortunately, access to those ideas and the results of research generated by them has not always been easy to obtain. Copies of the two key texts, *The Dynamic Assessment of Retarded Performers* (1979) and *Instrumental Enrichment* (1980) have been like gold-dust in this country, despite highly positive reviews when first published (e.g. reviews by Pumphrey in a DECP Newsletter). Moreover, those texts themselves are highly complex, theoretical presentations of quite revolutionary ideas which contain data from only the earliest implementation studies and their evaluations. Data from subsequent evaluations have tended to be reported at conferences or in funded research reports and not made widely available through professional journals.

It is mainly for these reasons that I produced a simple introduction to some of the main aspects of Instrumental Enrichment in the June, 1985, edition of *Remedial Education*, a slightly broader review and description of an Israeli workshop in *School Psychology*

*International*, 8 (1), 1987, and a critical review of most of the reported research studies in the *European Journal of the Psychology of Education*, 2 (1), 1987. One or two other review papers are also beginning to appear and more research reports are now finding their way into the literature. However, there has long been the need for a book with the aim of introducing Feuerstein and his ideas to the intelligent lay-person. What we all need from time to time is a summary of key ideas placed in a simple (but not simplistic) context which orientates us to further thinking and reading and which we can reliably recommend to others as a basic introduction to the area. I believe that with a few minor reservations Howard Sharron's *Changing Children's Minds* performs that task admirably.

Howard Sharron is basically a journalist and a good one. He writes in a no-nonsense style which captures and retains the reader's attention and has the happy knack of presenting complex psychological ideas in a straightforward, intellectually honest but jargon-free manner which does justice to his subject matter. As someone who has studied Feuerstein's theories in considerable depth over the past few years, I could only find one major fault with the simplified description offered in this book. The notion of the "cognitive map" is considerably more complex than is described.

What the book offers is an introductory background (all too brief) on Feuerstein the man and the roots of his theory, a brief description of the Learning Potential Assessment Device and Instrumental Enrichment, a summary of some of the more important and significant worldwide research and intervention studies, some tentative links with other theoretical perspectives such as those of Vygotskii and Peto, and a radical critique of the attitudes and actions of certain key figures in the processes of implementation and evaluation.

Be warned, however. Howard Sharron makes no claim to being objective. He wears his own ideology on his sleeve and makes no bones about the fact that this clashes with what he sees as established British psychological and educational thinking. For him Feuerstein offers the hope for a major breakthrough in ways in which such concepts as "intelligence" and "culture" are perceived and in action guidelines which are desperately needed in British education and society at the moment. I must say that I'm generally bound to agree with him. I also agree with his devastating indictment of the way in which Instrumental Enrichment was introduced into this country and subsequently "evaluated" by the Schools' Council. This is one of the clearest examples of how to administer the "kiss of death" on a potentially worthwhile intervention technique that I have ever come across.

I don't agree with Howard Sharron's rather outdated views on what would constitute a worthwhile evaluation design nor do I agree entirely with the "wart-free" picture that is painted here of the theory and research. There are problems with both which must be confronted. However, that's another story, as also is the fact that Feuerstein represents a team of workers, each of whom has contributed something of significance to the theory, practice and research. If you want a good starting point, then get hold of this book. You won't agree with all that it says, but you won't fail to be stimulated by it.

BOB BURDEN

---

VALSINER, J. (Ed.) (1986). *The Individual Subject and Scientific Psychology*. New York: Plenum Press, pp. 408, $45.

The editor of this book has brought together a group of psychologists who are concerned about the way psychology has tended to rely on statistical analysis of groups in building psychological theory. They believe that there is much to be learned from the detailed consideration of individuals and set out to demonstrate the value of such analyses through examples from a variety of contrasting specialisms.

The editor locates the origins of the book in an analysis of the types of study presented in child psychology journals between the 1930s and the 1980s. He found that, in spite of repeated warnings, there the number of studies investigating individual children had diminished continuously, to be replaced by studies in which the individual vanished beneath a welter of statements about means and "effects". His conclusion was that:

> "Empirical psychology, in its striving towards increased accumulation of data, may have actually eliminated the reality that remains behind the data . . . Psychology's major problem . . . has been the trivialisation of epistemological problems of research through

> quick and uncritical acceptance of different formalised systems of data analysis without careful scrutiny of their axiomatic background, and without a thorough evaluation of the fit between these methods with the important aspects of psychological phenomena that psychologists wish to study."

From that position this book sets out to show how data collected from individuals can be analysed in ways which are rigorous and yet retain information about the individual person's functioning in a variety of settings.

If the above quotation suggests a rather theoretical orientation that impression would be correct. The book is heavy going with its mixture of detailed theoretical discussions and descriptions of analytic procedures, and yet the issues raised are of great importance to educational psychology. The book does not provide a set of instructions about how to carry out research using case studies of individuals; rather it uses illustrations from several contrasting research areas to draw attention to the variety of approaches available to the researcher. The emphasis is on developmental psychology but with examples drawn also from social psychology, behaviour analysis, neuropsychology and psychophysiology.

There are only two chapters of direct relevance to educational psychology. Herbert Ginsburg from Teachers College, Columbia, discusses the importance of academic diagnosis and suggests the different approaches which have been used, historically, and the approaches which show the best potential for future research and applications. He provides a useful summary of some of the American work on the analysis of the cognitive processes used by children in carrying out everyday classroom tasks, but British research in this field is not mentioned. Perhaps the author should be directed towards the BJEP publication, *Recent Advances in Classroom Research.* While emphasising the importance of cognitive analysis and the work of Vygotsky and Feuerstein on "learning potential", Ginsburg is also ready to acknowledge the need to see the individual child in a broader perspective. Too often in the research using cognitive analysis we are left with the impression that we are dealing not with children but with a set of recalcitrant computers with "buggy algorithms". Ginsburg argues that we need to develop new integrative theoretical models directed towards the understanding of the whole child in the environment and directs the attention of researchers specifically to the need to explore the links between achievement motivation and cognitive processes, between locus of control and knowledge, and between social interaction and learning.

> "A comprehensive understanding of the child, one who is failing or excelling in schoolwork, must include consideration of general performance in school, attitudes and motivation, cognitive style, relations with parents and other adults, personality structure, and social environment. Such non-cognitive factors obviously play a major role in the child's academic performance. Thus, the understanding of motivation may be as least as useful for educational practice — for remediation — as knowledge of cognitive structure of process. Diagnosticians must and do perform integrative analyses all the time: in the real world, children's academic cognition is embedded in a larger psychological and social context."

The second chapter of interest to educational psychologists is written by Barbara Gogoff and Mary Gauvain. They discuss the analysis of patterns from mother-child instructional interactions. This chapter is concerned with establishing appropriate coding categories and possible problem-solving strategies by using graphic arrays and statistical procedures which focus on patterns of individual performance. These analyses also sought to combine the use of illustrative transcript material with quantitative estimates of the occurrence of patterns of interaction. This technique makes use of the sequences of acts and "the contextual relationships between the acts of different participants to develop a functional analysis of the workings of the event" and has potential applications in other areas of educational research. The findings in this chapter are of much less significance than the innovative analytic techniques adopted.

Overall this book contains a variety of stimulating ideas about the collection and analysis of data describing the behaviour of individuals and although the more theoretical chapters may be heavy going the discussions will be of considerable interest to researchers who share the editor's concern about the predominant approaches to research in child psychology.

NOEL ENTWISTLE

# PUBLICATIONS RECEIVED

Books for review have been selected from the following
("n.p." means "no price mentioned").

Alston, J., and Taylor, J. (1987). *Handwriting*. Beckenham: Croom Helm, pp. 229, £25·00 hbk., ISBN 0-7099-0580-7; £10·95 pbk., ISBN 0-7099-5107-8.

Altman, I., and Wandersman, A. (1987). *Neighborhood and Community Environments*. New York: Plenum, pp. 298, £39·50 hbk., ISBN 0-306-42513-0.

Beech, J., and Colley, A. (1987). *Cognitive Approaches to Reading*. Chichester: Wiley, pp. 315, hbk., ISBN 0-471-91169-0, n.p.

Betz, N. E. (1987). *Journal of Vocational Behaviour, Vol. 29, No. 3*. London: Academic Press, pp. 168, n.p.

Beveridge, M., and Conti-Ramsden, G. (1987). *Children with Language Difficulties*. Milton Keynes: Open University Press, pp. 116, £22·50 hbk., ISBN 0-335-10280-8; £6·95 pbk., ISBN 0-335-10279-4.

Blagg, N. (1987). *School Phobia and Its Treatment*. Beckenham: Croom Helm, pp. 228, £25·00 hbk., ISBN 0-7099-3938-8; £11·95 pbk., ISBN 0-7099-5050-0.

Brennan, W. K. (1987). *Changing Special Education Now*. Milton Keynes: Open University Press, pp. 173, £22·50 hbk., ISBN 0-335-10278-6; £7·95 pbk., ISBN 0-335-10277-8.

Brown, R. I., and Hughson, E. A. (1987). *Behavioural and Social Rehabilitation and Training*. Chichester: Wiley, pp. 192, £22·00 hbk., ISBN 0-471-91149-6.

Butler, D. (1987). *Cushla and her Books*. Harmondsworth: Penguin, pp. 123, £3·95 pbk., ISBN 0-14-009261-7.

Calam, R., and Franchi, C. (1987). *Child Abuse and its Consequences*. Cambridge: Cambridge University Press, pp. 226, £27·50 hbk., ISBN 0-521-30277-3; £7·95 pbk., ISBN 0-521-31612-6.

Clements, J. (1987). *Severe Learning Disability and Psychological Handicap*. Chichester: Wiley, pp. 186, hbk., ISBN 0-471-91354-5, n.p.

Colby, A., and Kohlberg, L. (1987). *The Measurement of Moral Judgment, Vol. 1*. (New York: Cambridge University Press) pp. 397, $54·50 hbk., ISBN 0-521-24447-1.

Coleman, J. (1987). *Working with Troubled Adolescents*. London: Academic Press, pp. 313, £20·00 hbk., ISBN 0-12-179720-1.

De Corte, E., Lodewijks, H., Parmentier, R., and Span, P. (1987). *Learning and Instruction*. Oxford: Pergamon, pp. 472, £30·00 hbk., ISBN 0-08-034769-X.

Digby, P. G. N., and Kempton, R. A. (1987). *Multivariate Analysis of Ecological Communities*. London: Chapman and Hall, pp. 206, £13·95 hbk., ISBN 0-412-24640-6.

Edwards, D., and Mercer, N. (1987). *Common Knowledge*, London: Methuen, pp. 193, £9·95 pbk., ISBN 0-416-39900-2.

Friedman, S. L., Scholnick, E. K., and Cocking, R. R. (1987). *Blueprint for Thinking*. Cambridge: Cambridge University Press, pp. 559, £40·00 hbk., ISBN 0-521-25605-4.

Fryer, D., and Ullah, P. (1987). *Unemployed People*. Milton Keynes: Open University Press, pp. 292, £25·00 hbk, ISBN 0-335-15507-3; £9·95 pbk., ISBN 0-335-15506-5.

Furth, H. G. (1987). *Knowledge as Desire*. New York: Columbia University Press, pp. 179, $22·50 hbk., ISBN 0-231-06458-6.

Gazda, G. M., Childers, W. C., and Brooks, D. K. (1987). *Foundations of Counsellling and Human Services*. London: McGraw Hill, pp. 327, £19·95 pbk., ISBN 0-07-022992-9.

Gottman, J. M., and Parker, J. G. (1987). *Conversations of Friends*. Cambridge: Cambridge University Press, pp. 429, £35·00 hbk., ISBN 0-521-26321-2.

Hannon, P., and Jackson, A. (1987). *The Belfield Reading Project Final Report*. London: National Children's Bureau, pp. 26, £2·95 pbk., ISBN 0-902817-33-7.

Hartup, W. W. (1987). *Child Development, Vol. 58, No. 4*. Chicago: University of Chicago Press, pp. 259, n.p.

Herbert, M. (1987). *Behavioural Treatment of Children with Problems*. London: Academic Press, pp. 276, hbk., ISBN 0-12-341470-9, n.p.

Hill, M. (1987). *Sharing Child Care in Early Parenthood*. London: Routledge and Kegan Paul, pp. 359, £29·50 hbk., ISBN 0-7102-0497-3.

Hodgson, V. E., Mann, S. J., and Snell, R. (1987). *Beyond Distance Learning Teaching — Towards Open Learning*. Milton Keynes: Open University Press, pp. 182, £25·00 hbk., ISBN 0-335-15603-7.

Holt, M., (1987). *Skills and Vocationalism*. Milton Keynes: Open University Press, pp. 179, £22·50 hbk., ISBN 0-335-10290-5; £7·95 pbk., ISBN 0-335-10289-1.

Honess, T., and Yardley, K. (1987). *Self and Identity.* London: Routledge and Kegan Paul, pp. 382, £25·00 hbk., ISBN 0-7102-0829-4.

Johnstone, M., and Munn, P. (1987). *Discipline in Schools.* Edinburgh: SCRE, pp. 50, £6·90 hbk, ISBN 0-947833-22-6, £3·90 pbk., ISBN 0-947833-20-X.

Jones, R. (1987). *Like Distant Relatives.* Aldershot: Gower, pp. 304, £25·00 hbk., ISBN 0-566-05323-3.

Laing, A. F., and Chazan, M. (1987). *Teachers' Strategies in Coping with Behaviour Difficulties in First Year Junior School Children.* Maidstone: AWMC, pp. 79, £4·00 pbk.

Langford, P. (1987). *Concept Development in the Primary School.* Beckenham: Croom Helm, pp. 133, £19·95 hbk., ISBN 0-7099-4162-5.

Lomas, P. (1987). *The Limits of Interpretation.* Middlesex: Penguin, pp. 158, £3·95 pbk., ISBN 0-14-022660-5.

Mitchell, S. (1987). *Nine Year Olds Grow Up.* London: Tavistock, pp. 181, £20·00 hbk., ISBN 0-422-78970-4.

Modgil, S., and Modgil, C. (1987). *Noam Chomsky.* Barcombe: Falmer Press, pp. 309, £35·00 hbk., ISBN 1-8500-022-0.

Orton, A. (1987). *Learning Mathematics.* London: Cassell Education, pp. 177, £7·95 pbk., ISBN 0-304-31369-6.

Parsons, M. J. (1987). *How We Understand Art.* Cambridge: Cambridge University Press, pp. 159, £19·50 hbk., ISBN 0-521-32949-3.

Reynolds, C., and Mann, L. (1987). *Encyclopedia of Special Education Vols. 1, 2 and 3.* Toronto: Wiley, pp. 1793, hbk., ISBN 0-471-82858-0, n.p.

Russell, J. (1987). *Philosophical Perspectives on Developmental Psychology.* Oxford: Basil Blackwell, pp. 262, £27·50 hbk., ISBN 0-631-14626-1.

Rutkowska, J. C., and Crook, C. (1987). *Computers, Cognition and Development.* Chichester: Wiley, pp. 311, hbk., ISBN 0-471-91583-1, n.p.

Sharron, H. (1987). *Changing Children's Minds.* London: Souvenir Press, pp. 336, £16·95 hbk., ISBN 0-2854-65034-3.

Sigmon, S. B. (1987). *Radical Analysis of Special Education.* Barcombe: Falmer Press, pp. 124, pbk., ISBN 1-85002-31-2, n.p.

Snowling, M. (1987). *Dyslexia.* Oxford: Basil Blackwell, pp. 162, £7·95 pbk., ISBN 0-631-14433-1.

Solity, J., and Bull, S. (1987). *Special Needs — Bridging the Curriculum Gap.* Milton Keynes: Open University Press, pp. 148, £22·50 hbk., ISBN 0-335-10282-4; £8·95 pbk.; ISBN 0-335-10281-6.

The Commission of the European Communities, Directorate General XIII (1987). *Development of European Learning through Technological Advance.* Paris: Les Editions du Logiciel d'Enseignement, pp. 626, pbk., ISBN 2-87625-001-2, n.p.

Vallender, I., and Fogelman, K. (1987). *Putting Children First.* Barcombe: Falmer Press, pp. 188, £8·50 pbk., ISBN 1-85000-219-3.

Wade, B., and Moore, M. (1987). *Special Children . . . Special Needs Provision in Ordinary Classrooms.* London: Robert Royce, pp. 219, £10·95 pbk., ISBN 0-947728-36-8.

Warr, P. (1987). *Psychology at Work.* Harmondsworth: Penguin, pp. 429, £6·95 pbk., ISBN 0-14-022654-0.

Weil, S. (1987). *The Need for Roots.* London: Routledge and Kegan Paul, pp. 288, £3·95 pbk., ISBN 0-7448-0058-7.

Weinstein, C. S., and David, T. G. (1987). *Spaces for Children.* New York: Plenum, pp. 322, £39·50 hbk., ISBN 0-306-42423-1.

Westwood, P. (1987). *Commonsense Methods for Children with Special Needs.* Beckenham: Croom Helm, pp. 156, £22·50 hbk., ISBN 0-7099-5210-4; £7·95 pbk.; ISBN 0-7099-5247-3.

## Clinical Attribution Research

A special issue of the *British Journal of Clinical Psychology*, guest edited by Chris R. Brewin.

Editorial: Developments in an attributional approach to clinical psychology—*Chris R Brewin*
The hopelessness theory of depression: Attributional aspects—*Lauren B. Alloy, Lyn Y. Abramson, Gerald I. Metalsky & Shirley Hartlage*
A longitudinal test of the attributional vulnerability model in children at risk for depression—*Constance Hammen, Cheri Adrian & Donald Hiroto*
Attributional change during psychotherapy—*Jenny Firth-Cozens & Chris R Brewin*
Attribution and adjustment to life-threatening illness—*Dawn C. Turnquist, John H Harvey & Barbara Anderson*
Causal attributions of doctor and patients in a diabetic clinic—*C. R. Gillespie & C Bradley*
The impact of attributions in marriage: Empirical and conceptual foundations—*Frank D. Fincham & Thomas N. Bradbury*
Causal beliefs amongst families in therapy: Attributions at the group level—*Anthony G Munton & Charles Antaki*
Attribution theory and attributional therapy: Some theoretical observations and suggestions—*Bernard Weiner*

Cognitive style and depression: Symptom-related, event-related or independent provoking factor?—*Glenys Parry & Chris Brewin*

The special issue will be available as Part 1 of the British Journal of Clinical Psychology, volume 27 (1988). Subscription to volume 27 (ISSN 0144-6657), edited by Professor Ray Hodgson, Whitchurch Hospital, Cardiff, is £49.50 (US$89.00).

Single copies of Part 1 only will be made available at the special price of £6.95 (US$11.95) for a limited period until June 1989 only.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

## Stress and Health

A special issue of the *British Journal of Medical Psychology*, guest edited by Cary L. Cooper and Sidney Crown.

Introduction—*Cary L. Cooper*
A structural model approach toward the development of a theory of the link between stress and mental health—*Cary L. Cooper & A. J. Baglioni, Jr*
The impact of social support on mental and physical health—*Daniel C. Ganster & Victor Bart*
Type A behaviour and the healthy individual—*John M. Ivancevich & Michael T. Matteson*
Immunological consequences of acute and chronic stressors: Mediating role of interpersonal relationships—*Susan Kennedy, Janice K. Kiecolt-Glaser & Ronald Glaser*
Environmental stress, reactivity and ischaemic heart disease—*David S Krantz & Shera E. Raisen*
Personality, stress and cancer: Prediction and prophylaxis—*H. J. Eysenck*

The special issue will be available as Part 1 of the *British Journal of Medical Psychology*, volume 61 (1988). Subscription to volume 61 (ISSN 007-1129), edited by Dr Sidney Crown, is £59.00 (US$106.00).

In order to make it available to the widest possible audience, the special issue is available as a single part, for a limited period only (until June 1989) at a special price of £6.95 (US$11.95) for the single journal part.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

## The Social Context of Emotion

A special issue of the *British Journal of Social Psychology*, guest edited by Hugh L Wagner.

Introduction—*Hugh L Wagner*
The social form of feeling—*Alan Radley*
Constructing emotion: Weaving meaning from memories—*Susan Kippax, June Crawford, Pam Benton, Una Gault & Jenny Noesjirwan*
Emotion/cognition interaction in personality development: A discrete emotions, functionalist analysis—*Carol Z Malatesta & Arnold Wilson*
Emotion and social context: An American–German comparison—*Shula Sommers & Corinne Kosmitzki*
Sex differences in interpretation of angry and weeping episodes—*Muriel Egerton*
Impression management and self-deception: Response bias components of appraised emotional experience—*Gordon L. Flett, Kirk R. Blankstein, Patricia Pliner & Carl Bator*
Moods and compliance—*Sandra Milberg & Margaret S. Clark*

The special issue will be available as Part 1 of the *British Journal of Social Psychology*, volume 27 (1988) Subscription to volume 27 (ISSN 0144-6665), edited by Dr A. S. R Manstead, is £49 50 (US$89.00)

In order to make it available to the widest possible audience, the special issue is available as a single part, for a limited period only (until June 1989) at a special price of £6.95 (US$11.95) for the single journal part.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

## Technological Change and Innovation

A special issue of the *Journal of Occupational Psychology*, guest edited by Chris Brotherton.

Strategic human resource management in the UK computer supply industry—*Paul R Sparrow & Andrew M. Pettigrew*
A longitudinal study of adjustment to technological change: From mass to computer-automated batch production—*Ann Majchrzak & John Cotton*
Parental influence and teenagers' motivation to train for technological jobs—*Glynis M. Breakwell, Chris Fife Schaw & John Devereux*
The development of computer applications by non-employed people in community settings—*Catherine Cassell, Mike Fitter, David Fryer & Leigh Smith*
Long-wave cycles of sociotechnical change and innovation: A macropsychological perspective—*Kenyon B De Greene*
Information technologies and organizations—lessons from the 1980s and issues for the 1990s—*Frank Blackler*
A history of social psychological reactions to new technology—*Jim Carliopo*
Robotics—A challenge for occupational psychology—*Oded Shenkar*

The special issue will be available as Part 1 of the *Journal of Occupational Psychology*, volume 61 (1988). Subscription to volume 61 (ISSN 0305-8107), edited by Dr David Guest, is £49.50 (US$89.00).

In order to make it available to the widest possible audience, the special issue is available as a single part, for a limited period only (until June 1989) at a special price of £6 95 (US$11.95) for the single journal part.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

*announcing . . .*

# NEW DEVELOPMENTS IN EDUCATIONAL ASSESSMENT

edited by
H. D. Black and W. B. Dockrell
*British Journal of Educational Psychology Monograph 3*

app 208 pages paperback app £7.50 ISBN 7073 0546 2
due for publication April 1988

CONTENTS
Editorial and 'It's Ma Faither That Hits Me'/Criterion-referenced Assessment: What Role for Research?/The Role of Descriptive and Normative Techniques in Examination Assessment/Grade-related Criteria/Issues in the Implementation of Grade Testing/The National Certificate/Accreditation of Work-based Learning/The Assessment of Competencies/Issues in the Assessment of Affective Characteristics and the Impact of Scottish National Surveys of Achievement on Policy and Practice/Research and Policy-making in Assessment: Some Unanswered Questions

***also available . . .***

Monograph 1 **BEHAVIOURISM AND LEARNING THEORY IN EDUCATION**
edited by D. Fontana (£7.50)

Monograph 2 **RECENT ADVANCES IN CLASSROOM RESEARCH**
edited by N. Bennett and C. Desforges (£7.50)

If you would like to purchase copies of these titles directly from the publisher please write to:

**Sales Department, Scottish Academic Press**
**33 Montgomery Street**
**Edinburgh EH7 5JX**

# CAMBRIDGE

## Were They Pushed or Did They Jump?

Individual Decision Mechanisms in Education

**DIEGO GAMBETTA**

This book explores the factors which govern the range of educational decisions confronting individuals between compulsory school education and university. Diego Gambetta explores the extent to which an individual's behaviour is 'pushed' by external constraints or casual mechanisms operating in the family background and to what extent 'pulled' by intentional and rational choices on the part of the individual.

*244 pp.* *0 521 32490 4* **£27.50 net**

*Studies in Rationality and Social Change*

## The Social Psychology of Education

Current Research and Theory

Edited by **ROBERT S. FELDMAN**

This important collection of papers examines the application of social psychology to education. The focus moves from the individual (e.g. student attitudes and attributions), to teacher-student interaction, to the impact of society (e.g. racial and cultural influences on school performance).

*381 pp.* *0 521 30620 5* **£35.00 net**

**Cambridge University Press**
The Edinburgh Building, Shaftesbury Road, Cambridge CB2 2RU, England.

## PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

*British Journal of Psychology*
*British Journal of Medical Psychology*
*Journal of Occupational Psychology*
*British Journal of Social Psychology*
*British Journal of Clinical Psychology*
*British Journal of Developmental Psychology*
*British Journal of Mathematical and Statistical Psychology*
*The British Journal of Educational Psychology*

The *British Journal of Psychology,* the *British Journal of Medical Psychology, Journal of Occupational Psychology,* the *British Journal of Development Psychology,* the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1988 subscription prices, payable in advance, are £74·50 (US$134·00) for the *British Journal of Psychology;* £49·50 (US$89·00) for the *British Journal of Social Psychology,* £49·50 (US$89·00) for the *British Journal of Clinical Psychology;* £45·00 (US$81·00) for the *British Journal of Developmental Psychology;* £59·00 (US$106·00) for the *British Journal of Medical Psychology;* and £49·50 (US$89·00) for the *Journal of Occupational Psychology.*

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £49·50 per volume post free (US$89·00 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts., SG6 1HN, UK. A special price of £340·00 (US$£610·00) for all seven journals going to the same address is offered for 1988.

*The British Journal of Educational Psychology* is issued three times a year, in February, June and November. The prepaid subscription is £20·00 UK, £22·00 overseas a year, post free, or £7·00 each number. The prepaid subscription price in the USA is $45·00 post free Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX

### MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D. E. Blackman, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Medical Psychology* should be sent to Dr Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.

For the *Journal of Occupational Psychology* should be sent to Dr. D. Guest, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Social Psychology* should be sent to Dr A. S R Manstead, Journals Office, 13A Church Lane, London, W2 8DX

For the *British Journal of Clinical Psychology* should be sent to Professor R Hodgson, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr A. D. Lovie, Journals Office, 13A Church Lane, London, W2 8DX

For the *British Journal of Developmental Psychology* should be sent to Professor George Butterworth, Journals Office, 13A Church Lane, London, W2 8DX.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Hazel Francis, University of London Institute of Education, 24-27 Woburn Square, London, WC1H 0AA.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR

**All manuscripts should be submitted in duplicate.**

### OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

### REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., PO Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

### SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £40·00

They are entitled to purchase any or all of the Society's Journals at the special rates of £7·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR

# BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 58, PART 1, 1988

## CONTENTS

P. E. Vernon: An Appreciation ........ i

ARTICLES

BUTLER, R. — Enhancing and Undermining Intrinsic Motivation: the Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance ........ 1

CLIFFORD, M. M. — Failure Tolerance and Academic Risk-Taking in Ten- to Twelve-Year-Old Students ........ 15

KIM, A., and CLIFFORD, M. M. — Goal Source, Goal Difficulty, and Individual Difference Variables as Predictors of Responses to Failure ........ 28

FOON, A. E. — The Relationship Between School Type and Adolescent Self Esteem, Attribution Styles, and Affiliation Needs: Implications for Educational Outcome ........ 44

SCHWARZWALD, J., HOFFMAN, M., and ROTEM, T. — A Preparatory Programme for Passage into the Ethnically Integrated School: Impact on School Climate and Social Relations ........ 55

ZEIDNER, M. — The Relative Severity of Common Classroom Management Strategies: The Student's Perspective ........ 69

WORRALL, N., WORRALL, C., and MELDRUM, C. — Children's Reciprocations of Teacher Evaluations ........ 78

CONSTABLE, H., CAMPBELL, R. M., and BROWN, R. — Sectional Drawings from Science Textbooks: an Experimental Investigation into Pupils' Understanding ........ 89

LAXON, V. J., COLTHEART, V., and KEATING, C. — Children Find Friendly Words Friendly too: Words with many Orthographic Neighbours are Easier to Read and Spell ........ 103

RESEARCH NOTES:

GREEN, L. F., and FRANCIS, J. — Children's Learning Skills at the Infant and Junior Stages: a Follow-On Study ........ 120

LISTER, C. M., LEACH, C., and WESENCRAFT, K. — Sequence in Hearing Impaired Children's Development of Concepts ........ 127

BOOK REVIEWS ........ 134

PUBLICATIONS RECEIVED ........ 141

UK ISSN 0007-0998

# THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

JUNE 1988

VOL. 58, PART 2

# THE
# BRITISH JOURNAL
# OF
# EDUCATIONAL PSYCHOLOGY

Edited by
Professor HAZEL FRANCIS

Assistant Editors:
Dr. H. C. M. CARROLL
Prof. C. W. DESFORGES
Dr. D. FONTANA
Dr. F. J. TAYLOR

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Prof. C. W. DESFORGES

**Advised by the Editorial Committee:**

Professor G. BROWN
Ms. I. LUNT
Dr. C. LEACH
Professor P. W. ROBINSON
Mr. E. WILKINSON

Vol. 58, Part 2

*The British Journal of Educational Psychology* is issued by the British Psychological Society once a term, in February, June and November. The subscription is £20.00 a year to the UK, £22.00 overseas, or £7.00 each number.

*Subscriptions and Orders* should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

*Papers for publication* in duplicate should be sent to Professor Hazel Francis, University of London Institute of Education, 24-27 Woburn Square, London, WC1H 0AA. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to: Prof. C. W. Desforges, University of Exeter, St. Luke's, Heavitree Road, Exeter, EX1 2LU.

*Advertisements* for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

*All rights reserved*

© 1988. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY LINDSAY & CO LTD , EDINBURGH

*Br. J. educ. Psychol.*, 58, 147-151, 1988

# CHILDREN'S PERCEPTION OF AND ANTICIPATED RESPONSES TO RISK

By GORDON C. A. DICKSON and GILLIAN E. HUTCHISON
*(Glasgow College)*

SUMMARY. The study examined children's awareness of, and potential response to, a number of different risks. These risks were categorised as either property risks (that is loss of or damage to property) or life risks (those causing personal injury to the child). Data were collected from 585 children aged 9 to 12 years old attending 12 primary schools in the Glasgow area. The tests were administered in booklet form during normal school hours. The results indicated that (a) the children were more aware of life risks than property risks; (b) the children responded more appropriately to potential property risks than to potential life risks; (c) the girls were more aware of life risks than the boys but there was no difference between boys and girls in terms of their behaviour towards potential life risks; (d) there was no difference between boys and girls in either their awareness of property risks or their anticipated behavioural response towards potential property risks.

## INTRODUCTION

WHILE the magnitude of risk can be measured in statistical terms, Sheppard (1975), Lichtenstein *et al.* (1978), Weinstein (1984) and others have shown that a conflict often emerges between the real risk and the risk as perceived by the individual. The conflict between real and perceived risk has then been used to explain, for example, apparently irrational behaviour (Lee, 1981), adoption of attitudes in social risks (Thomas, 1981), behaviour towards industrial risks (McLean, 1981) and the role of expert judgments (Slovic *et al.*, 1981).

With a few exceptions (Walesa (1977), Crader and Wentworth (1984) and Sheehy and Chapman (1986) being examples), the main emphasis of research has been on adults and their perception of risk. In particular researchers seem to have concentrated on how people process information about risk, rather than how they came to hold the views they do. The main focus of the research upon which this paper is based is children and the development of their thoughts about risk and uncertainty.

This paper reports one aspect of a large scale study carried out with children in the West of Scotland. The study is concerned with examining the extent to which children can recognise potential risk situations and the practical actions they say they would take in the face of risk. The long-term aim of the study is to make suggestions for classroom work based around the theme of educating children for risk and uncertainty.

## METHOD

### *Sample*

The sample comprised 585 children from 12 primary schools within the Glasgow Division of Strathclyde Regional Council. The sample was almost equally split between boys (289) and girls (296). All the children were in the final two years of their primary education and in the age range 9 to 12 (mean age 10·6 years).

The schools selected for the study represented a cross-section of areas within Glasgow Division. There were six schools in urban areas and six from suburban localities.

*Materials*

A test was devised which was aimed at measuring awareness of the factors which can cause risk and uncertainty. This was in booklet form and used language and situations which extensive pilot testing had suggested would be understood by children in the age range being studied. The booklet contained two different tests, each with its own objective. Test A was concerned with what action a child would take in the face of risk and Test B measured the extent to which children recognised situations to be risky.

*Test A*. This test comprised five potentially risky situations, three involving injury and two related to property. Each situation was followed by five possible actions which the child was asked to consider. These actions ranged from those which were judged to be the response of someone very aware of risk, to those of a person exhibiting little awareness. The actions had been judged previously by independent adult and child judges and each action had been assigned a number from 1 to 5, where a high number indicated a high level of awareness. The order in which the different actions appeared was varied for each situation.

To illustrate, as an example, one of the questions involved a story in which a child had taken his new bike to the shops and found he had forgotten his padlock. The five possible actions included: going home for the padlock, hiding the bike down an alleyway where few people would see it or asking someone to keep any eye on it while you are in the shop.

Earlier studies, such as those by Sieber (1969), have shown that children have a tendency to assume there is always a "correct" answer in any risky situation. To minimise the possibility of this occurring and in order to encourage a child to explore the range of alternative actions, each child was asked to mark the action or actions he definitely would not take. The score for each child was obtained by taking the rank of the action or actions chosen for each situation and summing for all five situations.

*Test B*. The second set of five questions, also involving three injury and two property situations, were of a different style. A statement was made and the child was asked to indicate how sure he was that the statement was accurate. The response was made by placing a cross on a straight line which had "very sure" at one end and "not sure" at the other. The statements related to potential risky situations such as, "Children who play with matches can start fires which could be dangerous to them and other people".

Scoring was achieved by placing a template over the line. This divided the line into five equal sections, each with its own numerical equivalent. The child's mark was then transformed to a score and the score for each of the five questions summed to produce the final score. A high score indicated a high level of recognition or awareness of risk.

These two tests were aimed at measuring different characteristics of the child's perception of risk. On the one hand there is awareness of risk and uncertainty and on the other there is behaviour when faced with risk. A child could be quite well aware of potential risk, but not take appropriate action. Thus Test B measured the extent to which the child recognised certain situations to be risky, while Test A looked at his potential behaviour in risky situations.

*Procedure*

The head teachers of the 12 schools in the sample were approached by the Education Authority and asked for their permission for the children to be used in the study. Each school was visited by a researcher who was introduced to the class by the class teacher. The teacher took no further part in the exercise. The researcher initiated a brief discussion on potential risky situations and possible preventative steps. The children offered a number of situations to which they felt exposed together with their suggested response. This led naturally to the introduction of the booklet. All children opened the booklet at the same time and were given whatever time was necessary to complete the questions. The researcher remained in the room while the children completed the booklet and offered assistance only to read aloud any question where necessary.

## RESULTS

TABLE 1

MEAN, MEDIAN AND MODAL SCORES, AND MEASURES OF SKEWNESS, FOR THE TWO DIFFERENT CATEGORIES IN TEST A AND TEST B

| | TEST A SCORE | | TEST B SCORE | |
|---|---|---|---|---|
| *Category* | *Property* | *Life* | *Property* | *Life* |
| Mean | 11·141 | 9·829 | 64·664 | 74·502 |
| Median | 11·559 | 9·915 | 66·25 | 80·693 |
| Mode | 13·000 | 9·000 | 81·5 | 81·333 |
| Skew | −1·479 | −0·443 | −1·070 | −2·523 |

The results indicated in Table 1 are for the two categories of question asked in Test A and Test B. In both tests the Life category comprised three questions concerning life threatening or injurious personal risks to the child, while the Property category comprised two questions concerning risks to the child's property.

The Life questions had a much lower mean, median and modal score in comparison to Property questions in Test A while in Test B the Life questions had higher mean and median scores and a very similar modal score to the Property questions. Of the measures of skewness for Test A the higher measure is for the Property questions. This indicated that for these questions children scored much more highly in terms of their potential behaviour when faced with risk to their property than they did when answering questions to do with their behaviour in the face of potential life threatening risks. This is even more surprising in the light of the measure of skewness for both categories in Test B. Here the opposite is the case and children scored far higher on the Life questions than they did on the Property questions when they were asked about their awareness of the risks involved in both categories.

Using the Wilcoxon test for differences between Property and Life scores in Test B, the latter were significantly higher ($P < 0{\cdot}001$), whilst, for Test A the reverse was true ($P < 0{\cdot}001$).

Similar results held for both boys and girls when their results were tested separately. The results for boys and girls were then compared within each of the two tests, using the Mann-Whitney test for statistical significance. In Test B (awareness) there was no significant difference in awareness of risk between boys and girls; however, the girls had substantially, if not significantly, higher scores in life risk situations ($P < 0{\cdot}07$). Further, there was no significant difference between boys and girls on either their potential behaviour towards property risks or towards life risks.

## DISCUSSION

One way of interpreting these results would be to say that the children used in the study appeared to be more aware of risks to their person than their property. However, when it came to suggesting how they would behave in the face of risk, their suggested behaviour seemed to indicate a much greater awareness of property risks. Why this should be so is much more difficult to explain than to report. The higher scores on awareness of life risks may be due, in part, to the increased level of publicity being given to assaults on children. The children in this study certainly recognised situations which were potentially risky as far as their personal well-being was concerned. However, the distribution of the scores for responding to such risks was much more symmetrical, indicating that the level of recognition was not matched by appropriate expected behaviour.

This pattern of scoring was not repeated for property risks. In these cases the indicated level of awareness was matched by their suggested behaviour. In other words, the children appeared to be aware of property risks, even if less so than life risks, and acted in a manner which also indicated risk awareness.

When the scores for boys and girls were looked at separately the results were much less conclusive but nevertheless interesting. There was no diference between the scores measuring response to risky situations, but the girls had a substantially higher score on awareness of life risks than did the boys. The result was just outside what would normally be considered statistically significant but large enough to suggest further enquiry.

What emerges from this study is the finding that the level of a child's expected response in the face of risk is not well predicted from his ability to recognise risk in the first place. Encouraging children to be aware of risky situations, as in the case of safety campaigns, is no guarantee that they will understand what kind of response is appropriate. Similarly it could be argued that training children to respond in specific ways in specific risky situations may simply increase the general level of awareness of risks covered by the training, but do little to increase the level of awareness for those risks not covered. In addition there is the clear indication of the need to separate risks into those involving personal injury and those related to property.

There can be no right answer as to how we should go about an educational programme for risk. This study suggests that children who learn how to make hypotheses about the risks which surround them, and who are helped to form appropriate expected actions from these hypotheses when faced with a risky situation, will have been better educated for facing risk.

Correspondence and requests for reprints should be addressed to Dr. G. C. A. Dickson, Department of Banking and Insurance, Glasgow College, Cowcaddens Road, Glasgow, G4 0BA.

## REFERENCES

CRADER, K. W., and WENTWORTH, W. M. (1984). A structural re-interpretation of responsibility, risk and helping in small collectives of children. *Am. Sociol. Rev.*, 49 (5), 611-619.

LEE, T. R. (1981). The public perception of risk and the question of irrationality. *Proc. R Soc. London*, Series A, Vol. 376, 5-16.

LICHTENSTEIN, S., SLOVIC, P., FISCHHOFF, B., LAYMAN, M., and COOMBS, B. (1978). Judged frequency of lethal events. *J. exp. Psychol.: Hum. Learn. Memory*, 4, 551-578.

MCLEAN, A. E. M. (1981). Assessment and evaluation of risks to health from chemicals. *Proc. R. Soc. London*, Series A, Vol. 376, 51-64.

SHEEHY, N. P., CHAPMAN, A. J. (1986). Adults and children's perceptions of hazard in familiar environments. In GARLING, T., and VALSINGER, J. (Eds.), *Children Within Environments*. London: Plenum Press.

SHEPPARD, D. (1975). The driving situations which worry motorists. Department of Environment. Transport and Road Research Laboratory, *Supplementary Report 129X*.

SIEBER, J. (1969). Lessons in uncertainty. *Elementary School J.*, 69, 304-312.

SLOVIC, P., FISCHHOFF, B., and LICHTENSTEIN, S. (1981). Perceived risk: psychological factors and social implications. *Proc. R. Soc. London.* Series A, Vol. 376, 17-34.

THOMAS, K. (1981) Comparative risk perception: how the public perceives the risks and benefits of energy systems. *Proc. R. Soc. London.* Series A, Vol. 376, 35-50.

WALESA, C. (1977). Development of risk perception in children and adults. *Polish Psychol. Bull.*, 8, 171-176.

WEINSTEIN, N. D. (1984). Why it won't happen to me: perceptions of risk factors and susceptibility. *Health Psychol.*, 3 (5), 431-457.

(*Manuscript received 12th June, 1987*)

**S. Sharan**, Tel-Aviv, **H. Shachar**, Tivon

**Language and Learning in the Cooperative Classroom**

1988 7 figures XVI, 176 pages
(Recent Research in Psychology)
Soft cover DM 54,-
ISBN 3-540-96708-7

**Language and Learning in the Cooperative Classroom** reports the results of an experiment on the effects of cooperative learning The authors address the following questions in detail Is the effect of cooperative learning on achievement more salient for pupils from the majority or minority ethnic groups? Do pupils who study with the Group-Investigation method display more extensive verbal interaction with their peers than pupils who study with the Whole-Class method? Is the pupil's verbal interaction in groups related to his/her academic achievement? The findings are discussed in terms of their relevance to classroom instruction, to different theories of language behavior and language research, and to issues regarding ethnic relations in multi-ethnic classrooms

**L. P. Steffe**, Athens, GA · **P. Cobb**, West Lafayette

**Construction of Arithmetical Meanings and Strategies**

With the collaboration of E. V Glasersfeld
Foreword by H Sinclair

1988 9 figures XIX, 348 pages (Recent Research in Psychology) Soft cover DM 76,- ISBN 3-540-96688-9

The studies presented in this book will be of interest to anybody concerned with the teaching of arithmetic to young children or with cognitive development in general The book provides an extremely detailed account of the different types of counting behavior of half a dozen children over two years The data allowed the authors to reach their main goal, to document the many subtle changes in children's counting and to interpret them theoretically At the same time, the authors cogently demonstrate that many of the widespread presuppositions about what young children know and what they do not know are erroneous, and that better insight into how children come to "do mathematics" should greatly improve the teaching of arithmetic

Springer-Verlag Berlin Heidelberg New York London Paris Tokyo

Heidelberger Platz 3, D-1000 Berlin 33 175 Fifth Ave , New York NY 10010 USA 28 Lurke Street, Bedford
MK40 3HU England 26 rue des Carmes F-75005 Paris 37-3, Hongo 3-chome, Bunkyo-ku Tokyo 113 Japan
Room 1603 Citicorp Centre 18 Whitfield Road, Causeway Bay, Hong Kong

Springer

b&s 4585/SF/2 q

*Br. J. educ. Psychol.*, 58, 152-164, 1988

# LEARNED HELPLESSNESS, SELF-WORTH MOTIVATION AND ATTRIBUTION RETRAINING FOR PRIMARY SCHOOL CHILDREN

By MARIE-LOUISE CRASKE
*(University of Tasmania, Australia)*

SUMMARY. This study examined characteristics of 29 pupils (selected from an initial sample of 69 primary school children) whose performance on an arithmetic task deteriorated after failure. On the basis of their responses to a "mitigating circumstance" which could explain failure without implicating low ability as the cause (a description of the task as "very difficult"), subjects were classified as either learned helpless or motivated to protect self-worth. Both of these groups had lower self-concepts than the rest of the initial sample. Attribution retraining resulted in increased effort attributions and decreased inability attributions in the "learned helpless" group, and inoculated these subjects to the experience of failure. In the "self-worth group" there was an increase in effort attributions but no change in ability ratings or performance after failure, following training.

## INTRODUCTION

PAST research has established that repeated failure can disrupt academic performance, resulting in decreased persistence and achievement levels (e.g., Dweck and Reppucci, 1973; Dweck and Bush, 1976; Andrews and Debus, 1978; Diener and Dweck, 1980). Two possible explanations for this effect are found in the reformulated learned helplessness model (Abramson *et al.*, 1978) and the self-worth theory of achievement motivation (Beery, 1975).

According to the former, a state of learned helplessness is reached when an individual perceives that he lacks control in obtaining a desired outcome. The type of explanation (attribution) the individual makes for his lack of control determines the features of his helplessness. For example, an internal, stable and global attribution will result in depressed affect, diminished self-esteem, low expectancy for future success and deteriorated performance. Thus, in an academic context, a student who has failed repeatedly at a particular task, and who construes the failures as a consequence of his lack of ability, will experience negative affect and a lowering of his self-esteem, and he will not expect to perform well on a similar task in the future. In particular, he will perform more poorly after failure than before, on tasks of equal difficulty (Dweck and Reppucci, 1973; Diener and Dweck, 1978).

Children with poor academic self-concept appear to be particularly susceptible to learned helplessness (Butkowsky and Willows, 1980). It is possible that a self-reinforcing cycle exists with a low self-concept predisposing attributions to lack of ability which then mediate reduced persistence and attainment levels. These in turn serve to maintain the self-concept at a low level. Females are more likely than males to exhibit learned helplessness (e.g., Dweck and Gilliard, 1975; Le Unes *et al.*, 1980; Wilson *et al.*, 1980).

In essence, the learned helplessness model implies that some students may "give up trying" because they do not see themselves as capable of success. Whether or not effort is applied, the outcome will be the same: failure. Logically, there is little to be gained by trying, and nothing to be lost by not trying.

In contrast, the self-worth theory states that, in some circumstances, students stand to gain a great deal by not trying. The self-worth theory is based on the notion that much of a student's behaviour is designed to maintain a self-concept of high ability. To this end, it is important to avoid failure whenever possible since failure carries with it implications of low ability. On the occasions when failure is unavoidable, low ability inferences can be deflected by ascribing it to stable, external factors (such as task difficulty) or to unstable elements (such as bad luck and insufficient effort). From this perspective, it is clear that the application of effort under conditions of possible failure can be risky: if a student tries hard but fails, then suspicions of low ability are increased. A reduction in, or withdrawal of, effort after a failure experience can therefore be used by the student as a strategy to prevent further damage to his sense of self-worth. Consequences of the withdrawal of effort are decreased persistence and achievement levels.

It has been noted that males are more inclined to defend their self-worth than females (Snyder *et al.*, 1978; Covington and Omelich, 1979). They also prefer to obtain successes through high ability, rather than the application of effort (Covington and Omelich, 1981). The relationship between self-concept and self-worth motivation has yet to be precisely determined. Evidence that males have higher self-concepts than females (Ickes and Layden, 1978), and that they are also more likely to defend their self-worth, suggests a tendency for high self-concept individuals to prefer self-serving attributions, and this has been reported in at least one study (Schwarzer and Jerusalem, 1982). However, students low in academic self-concept have also been reported as high in self-worth motivation (Covington and Omelich, 1979).

These two models represent different explanations for persistence and attainment decrements after failure. Two attempts have been made to determine which model provides the "better" explanation. Frankel and Snyder (1978) and Snyder *et al.* (1981) found that college students produced improved performances after failure under conditions (either high task difficulty or distracting background music) which could provide an explanation for failure without implying low ability. Learned helplessness predicts that such conditions would increase the subjects' perception of lack of control and therefore lead to worsened performances. This did not occur, and the authors interpreted the results as support for the self-worth (or egotism) model. However, the issue as to which is the most appropriate explanation is far from settled. A major difficulty in making comparisons has been the different subject populations used: primary school children have been the focus of learned helplessness research, and college students the target of self-worth investigations.

Numerous studies have established attribution retraining as an effective means of overcoming or inoculating against helplessness (e.g., Dweck, 1975; Andrews and Debus, 1978; Fowler and Peterson, 1981; Medway and Venino, 1982). Attribution retraining is designed to teach students to construe lack of effort as the cause of failures, and to recognise that success is more likely with effort. Craske (1985) used vicarious attribution retraining with "helpless" subjects who were selected from a larger population on the basis of deteriorated performances after failure. It proved to be an effective intervention for female subjects, but not for males. One possible explanation for this result assumes that some of the subjects, particularly males, were not "helpless" but had reduced effort after failure in an attempt to protect self-worth. Therefore, they did not respond to a technique designed to overcome "helplessness".

It seems logical to propose that both self-worth protection and learned helplessness exist amongst primary schoolchildren. It is proposed that some children decrease persistence after failure because they wish to avoid the appearance of inability, and others because they believe they are unable to control outcomes and trying is therefore irrelevant. Distinguishing between these two populations is important, since attempts to overcome low persistence will have varying effects depending on the reason for its existence. While it may be useful to train "helpless" students to attribute failure to low effort, to do so with self-worth motivated students may decrease persistence further. (If failure still occurs, even with the application of more effort, inability is strongly implicated.)

The main purpose of this study is to distinguish between those children whose deteriorated performance after failure is a consequence of learned helplessness, and those whose performance is motivated by self-worth considerations. It is predicted that these two groups will differ in their attributions for failure, and in their response to an attribution retraining procedure. It is also predicted that there will be more males in the latter category and more females in the former. In addition, the relationship of self-concept to learned helplessness and self-worth motivation will be investigated.

## METHOD

### *Sample*

Thirty-five male and 34 female pupils from a state primary school near Hobart constituted the initial sample for the study. They were drawn from Grade 4, 5 and 6 classrooms. The mean age was 10 years 7 months for both males and females. Grades 4 and 6 were taught by females, and Grade 5 by a male teacher. The catchment area of the school is mixed in terms of socio-economic status but is predominantly lower middle class.

### *Pretraining measures*

Pretraining assessment consisted of measures of self-concept, performance after failure, attributions for failure, and response to a "mitigating circumstance" which could account for failure without implicating low ability.

*Self-concept.* This was assessed using the School Form of the Coopersmith Self-Esteem Inventory (Coopersmith, 1981). The SEI is a 58-item questionnaire designed to measure evaluative attitudes toward the self in social, academic, family and personal areas of experience. A Lie Scale is also included. Kimball (1973) obtained reliability coefficients ranging from 0·87 to 0·92 for Grades 4 to 8; Spatz and Johnston (1973) obtained a coefficient of 0·81 for Grade 5. Adequate concurrent and construct validity have been reported by Kokenes (1974, 1978), and Simon and Simon (1975).

Pupils completed the inventory in classroom groups. None recorded extreme Lie scores.

*Performance after failure and response to a "mitigating circumstance".* These were assessed using four sets of sums: Sets A-D. The sets consisted of 10 sums for Grade 4 pupils, 15 sums for Grade 5 pupils, and 18 sums for Grade 6 pupils. Sets A, C and D were constructed so as to be within the capability of each child, according to his/her teacher's report, and consisted of sums requiring the operation of the four basic processes: addition, subtraction, multiplication and division. The sums were classified according to a number of categories (with and without regrouping, short and long division, short and long multiplication, inclusion of a zero in the figure, decimal figures). These categories were represented consistently across Sets A, C and D, although the actual numbers were altered. In this way the sets were matched for expected difficulty. Children were told that they were required only to do as many sums as they wished.

An addition was made to the instructions for Set D: "The sums in Set D are much harder than the others you have just done. I do not expect you to get many right. Just do as many sums as you like. You can stop when you want to." This instruction was included to provide an explanation for failure which does not implicate low ability. Set B was designed to provide a failure experience. It consisted of sums beyond the capability of the pupils. They were required to complete all of the sums on this set, and the salience of failure was enhanced by the experimenter marking the sums in the child's presence. At least two-thirds of the sums were marked as incorrect. Following the marking of Set B, pupils were required to indicate whether they had made some mistakes or achieved correct answers on all the sums, and to rate how well they felt they had done on Set B (very poorly, just OK, or very well). These questions were included to provide some check on the success of the experimental manipulation, i.e., that pupils did perceive Set B to be a "failure experience". Answers to the questions indicate that this was the case.

Attainment on Sets A, C and D was measured by a composite score taking into account both the number of sums attempted and the number correct in each set.

The index of performance after failure was calculated by subtracting the attainment score on Set C from the attainment score on Set A.

*Attributions for failure.* Scales to measure attributions for failure were included following Set B. Attributions used were the four commonly employed by students to explain their performances, and described by Weiner (1979): luck, task difficulty, effort and ability. Pupils were asked to indicate on a 7-point scale (ranging from 1: not at all important, to 7: very, very important) how important bad luck, the difficulty of the sums, lack of trying, and lack of ability were as causes of their mistakes. The Grade 4 children were given prior instructions in the use of such scales by the experimenter, and no difficulties were apparent.

Following the procedures used by Arkin and Maruyama (1979), and Gollwitzer *et al.* (1982), the four attributional measures were combined to create two distinct dimensions: internal-external and stable-unstable. These dimensions are delineated in Weiner's taxonomy of causes (1972, 1974). An index of internality was obtained by subtracting luck plus difficulty from the sum of effort plus ability scores, and of stability by subtracting luck plus effort from the sum of ability plus difficulty scores. The possible range for these composite scores is $-12$ to $+12$, with positive scores indicating internal and stable attributions and negative scores indicating external and unstable attributions.

### *Selection of Learned Helpless (LH) and Self-Worth (SW) groups*

Pupils whose performance deteriorated after failure (i.e., whose score on Set C was lower than the score on Set A) were selected for training in the second phase of the study. Twenty-nine (42 per cent of the initial sample) fell into this category.

Division into LH and SW groups was based on their responses to the "mitigating circumstance". Those who did better in Set D than A were classified as subjects motivated by self-worth considerations, i.e., their performance improved when the necessity for protecting self-worth was removed. Eleven pupils (five girls and six boys) fell into this category. Those who did better in Set A than D were considered to be displaying learned helplessness: their performance had already been disrupted by failure and the description of high task difficulty produced a further deterioration. Eighteen pupils (10 girls and 8 boys) were classified in the LH group.

### *Training*

Training was based on the reattribution approach; it was designed to counteract the disruptive effect of failure by encouraging pupils to view failures as the result of lack of effort, and to view subsequent successes as more likely with increased effort.

Pupils worked together in groups of five or six same-sex children, whose capability at sums was roughly equivalent. Thirty sums were displayed on a blackboard. The sums were arranged such that one in three was extremely difficult (and failure was anticipated on these) and no more than two "failure" sums occurred consecutively. The remaining sums were within the pupils' capabilities, and success was expected.

Pupils were told that they had been selected to play a maths game, with the following procedure. Each person in rotation had the opportunity to throw two dice. Two sixes earned a chance to work some sums on the blackboard, as many as could be completed until the next double six was thrown by another member of the group. When this occurred, the next person came to the board and began to work a new sum. Each time a sum was completed the game playing was stopped, and the correctness or otherwise of the solution was signalled by a green light for a right answer, and a red light for a wrong answer. At this point, the pupil who had completed the sum was asked to make an attribution for the outcome by selecting from the following choices which were displayed on a large chart:

| | |
|---|---|
| I got the sum right because: | I had good luck |
| | It was easy |
| | I tried hard |
| | I am clever |
| I got the sum wrong because: | I had bad luck |
| | It was too hard |
| | I didn't try hard enough |
| | I'm not clever enough |

Spontaneously made effort attributions were verbally reinforced by the experimenter (e.g., "Yes, that's right. You got the sum right because you did try hard"). Minimal cueing was allowable when spontaneous effort attributions were not forthcoming (e.g., "I think you got the sum wrong because you weren't trying hard enough"). In practice such cues were rarely needed, other than in the first few minutes of the game which lasted between 30 to 40 minutes in all.

The fact that children were required to make attributions publicly may have influenced their attributions, though any influence is likely to have been in the desired direction (House, 1980).

The procedure aimed to maximise cost-effectiveness (by working with groups rather than individuals) and the benefits of participant modelling. Previous research (e.g., Brown and Inouye, 1978; De Vellis *et al.*, 1978; Breen *et al.*, 1979; Zimmerman and Blotner, 1979; Zimmerman and Ringle, 1981; Craske, 1985) has indicated the usefulness of vicarious learning in reducing the disruptive effect of failure. Although it is possible that presenting training in the form of a dice game may have reduced its significance to some children, it was decided to do so in order to reduce the possible inhibiting effect of self-consciousness. Brockner (1979) suggests that the performance of low self-esteem individuals can be improved by reducing anxiety before making them self-aware.

### *Post-training measures*

The second phase of the study was designed to measure the effectiveness of training on the disruptive effect of failure, and on attributions for failure. Post-training assessment consisted of administration of Sets E, F and G. These were matched for expected difficulty with Sets A, B and C and followed the same procedure and instructions. As before an index of performance after failure was obtained by subtracting the attainment score on Set G from the attainment score on

Set E. Attributions for failure were assessed as before, following completion of Set F (the failure experience). These measures were taken on the same day or the day following training.

All pupils were debriefed after post-training measures were completed, and emphasis was placed on the fact that failures in Set B were contrived by the experimenter, and therefore did not reflect on pupils' ability levels. Many pupils expressed surprise that the outcome had been manipulated in this way.

## RESULTS

The main purpose of this study was to examine the self-concept, attributions and response to attribution retraining of those children whose performance deteriorated after failure. Analyses were also conducted on the scores of the initial sample of 69 children. Although the dependent variable measures probably do not constitute interval scales, the use of parametric statistical tests is still appropriate provided the score distributions do not markedly depart from the normal form (McNemar, 1969). Accordingly parametric procedures were used in analysing the results.

Table 1 shows the mean scores obtained from the initial sample at pre-testing.

### *Initial sample*

No significant sex differences were noted: boys and girls did not differ in terms of self-concept, the disruptive effect of failure, attributions for failure, or response to a mitigating circumstance.

Relationships between self-concept and A-C (response to failure), and self-concept and A-D (response to a mitigating circumstance) were examined. Because Sets A, C and D varied in length for Grades 4, 5 and 6, a pooled within-groups correlation (Keppel, 1982) was obtained, rather than a direct correlation. The

TABLE 1

Means and Standard Deviations of Scores on Pretraining Measures for Initial Sample

| | *Boys* (*N* = 35) | | *Girls* (*N* = 34) | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| Self Concept | 33·11 | 7·03 | 31·29 | 7·81 |
| Performance after Failure[a] | 0·80 | 3·72 | 0·41 | 3·46 |
| Attributions for Failure[b] | | | | |
| Luck | 3·91 | 1·87 | 4·82 | 1·99 |
| Difficulty | 3·86 | 1·77 | 4·15 | 1·67 |
| Effort | 4·26 | 1·90 | 4·77 | 1·72 |
| Ability | 4·23 | 1·75 | 4·79 | 1·87 |
| Composite Scores[c] | | | | |
| Internality | 0·71 | 2·07 | 0·59 | 3·06 |
| Stability | −0·03 | 4·07 | −0·56 | 3·68 |
| Response to a Mitigating Circumstance[d] | 0·17 | 4·47 | 0·91 | 4·36 |

a Difference between attainment score on Set A and attainment score on Set C.
b Range = 1 (not important at all) to 7 (very, very important).
c Range = −12 (external, unstable) to +12 (internal, stable).
d Difference between attainment score on Set A and attainment score on Set D.

pooled correlation between self-concept and A-C was significant ($r = 0\cdot30$, $P < 0\cdot05$) i.e., failure was most disruptive to the performance of low self-concept children. Self-concept was not related to response to a mitigating circumstance ($r = 0\cdot09$, NS).

Contrary to expectation, no significant correlations were found between self-concept and attributions for failure.

*LH and SW groups*

The 29 pupils whose performance deteriorated after failure were selected for training. Means obtained at pre-testing appear in Table 2.

These pupils had significantly lower self-concepts ($t\ (67) = 2\cdot96$, $P < 0\cdot01$), and, inconsistently, rated both stable factors ($t\ (67) = -3\cdot35$, $P < 0\cdot01$) and lack of effort ($t\ (67) = 3\cdot69$, $P < 0\cdot001$) as more important causes of their failures than the remainder of the sample. As a group they responded to being informed of increased task difficulty more poorly than other subjects (i.e., their A-D scores were larger) ($t\ (67) = -3\cdot15$, $P < 0\cdot05$).

TABLE 2

MEANS AND STANDARD DEVIATIONS OF SCORES ON PRETRAINING MEASURES FOR LH AND SW SUBJECTS

| | | *LH* | | *SW* | |
|---|---|---|---|---|---|
| | | *Boys (N=8)* | *Girls (N=10)* | *Boys (N=6)* | *Girls (N=5)* |
| Self-Concept | M | 32·13 | 28·20 | 29·17 | 27·00 |
| | SD | 4·83 | 7·35 | 4·41 | 3·35 |
| Performance after Failure[a] | M | 4·63 | 4·00 | 4·00 | 2·20 |
| | SD | 2·29 | 2·72 | 5·65 | 1·47 |
| Attributions for Failure[b] | | | | | |
| Luck | M | 3·00 | 4·70 | 3·67 | 5·40 |
| | SD | 2·18 | 1·36 | 1·49 | 1·63 |
| Difficulty | M | 4·25 | 4·00 | 4·50 | 3·80 |
| | SD | 1·71 | 1·43 | 2·06 | 0·40 |
| Effort | M | 3·25 | 4·50 | 3·67 | 4·80 |
| | SD | 1·71 | 1·61 | 1·60 | 1·60 |
| Ability | M | 6·00 | 6·10 | 4·17 | 4·40 |
| | SD | 0·87 | 1·48 | 1·77 | 1·02 |
| Composite Scores[c] | | | | | |
| Internality | M | 2·00 | 1·90 | 0·33 | 0·00 |
| | SD | 1·66 | 2·47 | 2·43 | 2·10 |
| Stability | M | 4·00 | 1·10 | 1·33 | −2·00 |
| | SD | 4·09 | 3·52 | 4·11 | 2·28 |
| Response to a Mitigating Circumstance[d] | M | 4·75 | 5·40 | −2·83 | −1·20 |
| | SD | 4·55 | 3·58 | 2·85 | 0·40 |

a Difference between attainment score on Set A and attainment score on Set C.
b Range = 1 (not important at all) to 7 (very, very important).
c Range = −12 (external, unstable) to +12 (internal, stable).
d Difference between attainment score on Set A and attainment score on Set D.

TABLE 3
MEANS AND STANDARD DEVIATIONS OF SCORES ON POST-TRAINING MEASURES

| | | *LH* | | *SW* | |
|---|---|---|---|---|---|
| | | *Boys (N=8)* | *Girls (N=10)* | *Boys (N=6)* | *Girls (N=5)* |
| Performance after Failure[a] | M | −1·50 | −1·70 | 2·67 | 3·20 |
| | SD | 1·00 | 3·41 | 7·43 | 3·97 |
| Attributions for Failure[b] | | | | | |
| Luck | M | 3·50 | 4·90 | 5·67 | 4·20 |
| | SD | 2·35 | 1·81 | 0·94 | 1·94 |
| Difficulty | M | 4·25 | 4·50 | 5·17 | 4·40 |
| | SD | 2·05 | 1·50 | 1·07 | 1·62 |
| Effort | M | 5·88 | 6·00 | 6·17 | 5·60 |
| | SD | 1·97 | 1·18 | 0·69 | 1·74 |
| Ability | M | 3·88 | 5·30 | 4·50 | 5·80 |
| | SD | 1·90 | 1·27 | 1·50 | 1·17 |
| Composite Scores[c] | | | | | |
| Internality | M | 2·00 | 2·90 | −0·17 | 2·80 |
| | SD | 2·40 | 3·73 | 1·34 | 2·04 |
| Stability | M | −1·25 | −1·10 | −2·17 | 0·40 |
| | SD | 3·03 | 1·22 | 2·12 | 2·33 |

a Difference between attainment score on Set E and attainment score on Set G.
b Range = 1 (not important at all) to 7 (very, very important).
c Range = −12 (external, unstable) to +12 (internal, stable).

Of most interest is the differentiation between LH and SW groups. By definition, the LH group performed much more poorly than the SW group when told of increased difficulty; it also tended to rate lack of ability more highly (F (1,25) = 9·98, P < 0·01). This was reflected in higher internality scores (F (1,25) = 5·27, P < 0·05). LH and SW groups did not differ from each other in self-concept (F (1,25) = 0·78, P > 0·05) or in the extent to which failure disrupted performance, (F 1,25) = 1·11, P > 0·05).

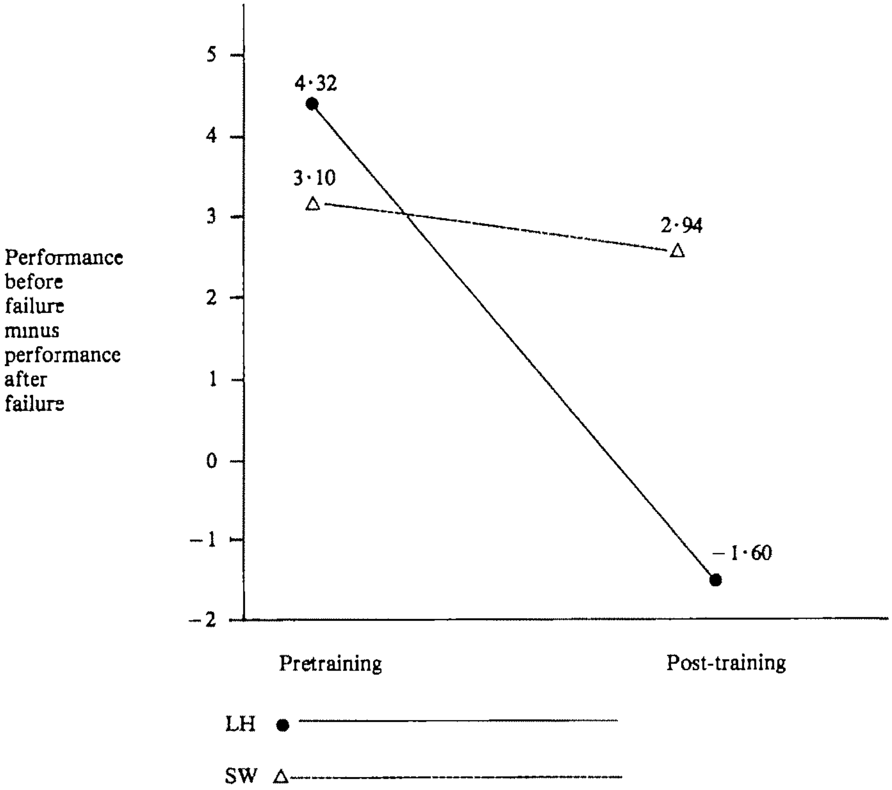
*Effect of training*

In the second phase of the study, the central hypothesis concerned the effectiveness of reattribution training in reducing the disruptive effect of failure. It was predicted that only the LH group would respond to training. Interest also lay in the effect of training on attributions for failure. Table 3 shows the mean scores obtained for post-training variables.

The effect of training on decrement in performance following failure (E-G scores) was examined using analysis of covariance, to make allowance for differences in set lengths as reflected in pre-training (A-C) scores. The assumption of homogeneity of regression (Keppel, 1982) was tested. No evidence of non-homogeneity was found (F 3,21) = 0·26, NS).

The covariance analysis showed a significant main effect for group (F (1,24) = 6·38, P < 0·05). After training, the effect of failure was significantly less disruptive for the LH group than it was for the SW group (see Figure 1). Thirteen of the LH group (N = 18) actually improved after failure at post-testing, compared to two of the SW group (N = 11).

FIGURE 1
THE EFFECT OF TRAINING ON PERFORMANCE AFTER FAILURE

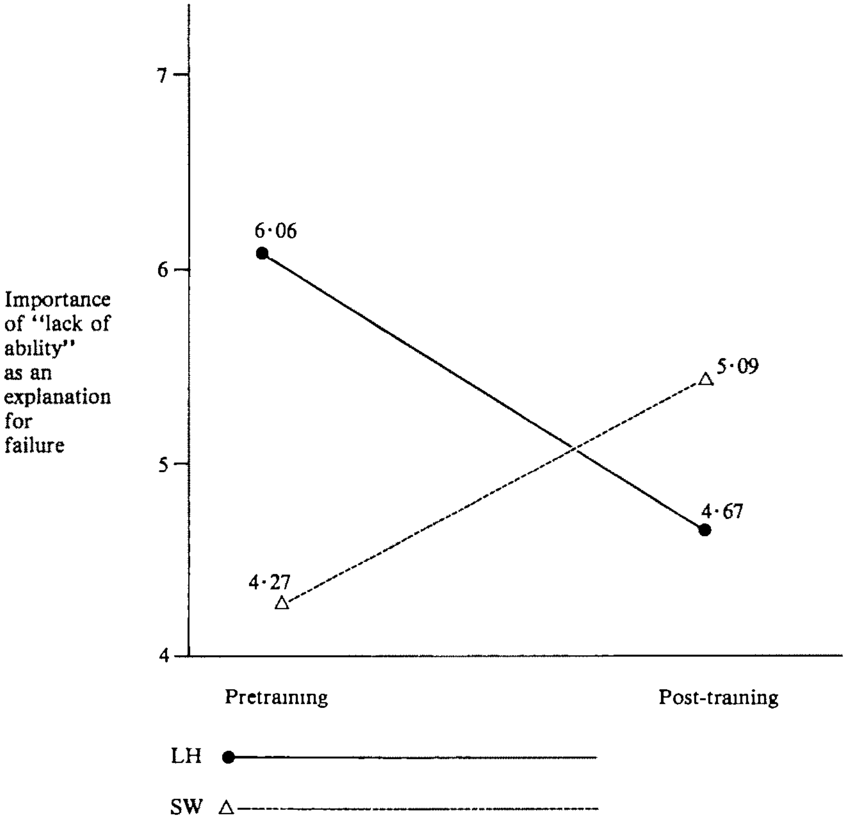


Two factor analyses of variance (Groups × Occasions) were conducted to evaluate the effect of training on attributions. Significant main effects for Occasions were noted on the effort (F (1,27) = 18·05, P < 0·001) and stability (F (1,27) = 4·71, P < 0·05) factors. After training lack of effort was rated more highly than before, and this change was reflected in lower stability scores.

A significant main effect for groups was found on the internality factor (F (1,27) = 5·36, P < 0·05). The LH group had consistently higher internality scores than the SW group at both pre- and post-testing.

Of most importance is the Groups × Occasions interaction effect, which determines whether attributional differences between LH and SW groups altered as a result of training. This was found to be significant for the ability factor (F (1,27) = 9·31, P < 0·01). Training influenced the ability ratings of the two groups differently. At post-testing the LH group placed less importance on lack of ability as a cause of failure, while the SW group tended to rate lack of ability slightly more highly (see Figure 2).

FIGURE 2
THE EFFECT OF TRAINING ON ABILITY ATTRIBUTIONS



## DISCUSSION

This study was designed to examine differences between those pupils whose performance deteriorated after failure because they felt academically helpless, and those who were protecting a sense of self-worth. Of most significance was the differing response of these two groups to the attribution retraining procedure. The LH group appeared inoculated to the experience of failure (or even motivated by it), while the SW group did not respond to training. This finding is in line with other research (e.g., Dweck, 1975; Andrews and Debus, 1978; Fowler and Peterson, 1981; Medway and Venino, 1982; Craske, 1985), which has also found attribution retraining to be an effective means of improving performance after failure, for pupils in the LH category.

The intervention also produced attributional change. After training, both groups increased the emphasis placed on lack of effort as a cause of failure. This was expected since the intervention encouraged subjects to explain their performances in terms of the application of effort. Similar results have been found by Dweck (1975), Andrews and Debus (1978), Gatting-Stiller *et al.* (1979) and Fowler and Peterson (1981).

Attribution retraining is based on the rationale that increased emphasis on effort mediates improved performances. It is assumed that as a child comes to understand that his mistakes occur because he has not applied enough effort, and that he can improve his performance by trying harder, his perceptions of inability will decrease, and there will be a corresponding increase in feelings of self-efficacy and expectations for future success. Such expectations are crucial determinants of subsequent, improved performances (Bandura, 1981; Schunk, 1983, 1984).

This process appears to be operating in the LH group and is reflected in their increased effort ratings, decreased ability ratings and better performances in the face of failure, after training. Although no post-training measure of self-concept was taken, an increase in self-concept would be predicted. Such an increase is also suggested by the correlation found in the initial sample between self-concept and A-C scores, i.e., children with higher self-concepts found failure to be less disruptive to their performance than their peers with lower self-concepts.

Pupils in the SW group did not show the same responses to training. They learnt to emphasise effort, but did not make the corresponding changes in ability ratings or performance. The reason for this is not clear. However, the level of pretraining ability ratings may be important. Prior to the intervention, the SW group placed significantly less importance on lack of ability as a cause of failure than did the LH.

Alternatively, it is possible that the improved performance exhibited by the LH group was unrelated to attributional change, but occurred because the experimenters' encouragement of attributions to lack of effort was perceived as an instruction to try harder. Hence, the LH group's increased persistence following intervention could simply reflect a compliance with the instruction. If this is the case, then the SW group's fear of "trying and failing" may have prevented them from complying in the same way that the LH group did. Covington and Omelich (1979) address this issue in describing effort as "the double edged sword in school achievement". When teachers give instructions to try hard and when they reward effort they may be presenting some students with a difficult choice: between compliance (and praise), on the one hand, and self-worth protection, on the other.

It is clear that further work is required to clarify the precipitants of behaviour change, its durability and generalisability, and the exact links between behaviour and attributions. Weiner's theory of achievement motivation does not postulate a direct link between causal attributions and behaviours, but includes additional intervening variables such as affects and expectancies. These intervening variables were not assessed in this study and have not received the same amount of research attention as attribution.

In addition, differences between LH and SW groups need to be examined further. From this study, both groups appear to have a number of features in common — they are not differentiated by sex or the extent to which they find failure to be disruptive. Both groups have lower self-concepts than their peers, but not from one another. Yet they had markedly different responses to the intervention. Attribution retraining (for whatever reasons) was effective in improving the performance of LH children, but appropriate interventions for SW children have yet to be investigated. Beery (1975) suggests that any treatment for the latter group must be based on the development of a sense of self-worth which is not contingent on ability.

The lack of sex differences on any variables is worthy of a final comment. A large amount of research has shown that girls are more likely to exhibit LH (e.g., Crandall, 1969; Dweck and Reppucci, 1973; Dweck and Gilliard, 1975; Gody, 1978;

Wilson *et al.*, 1980; Le Unes *et al.*, 1980) and to attribute failure to a lack of ability (e.g., Dweck and Reppucci, 1973; Nicholls, 1975, 1979; Dweck and Bush, 1976; Dweck *et al.*, 1978; Craske, 1985), while boys are more likely to be motivated by self-worth considerations (Covington and Omelich, 1979; Snyder *et al.*, 1978). As Craske (1985) notes, the influence of a male teacher may be significant. It has been argued that male teachers are less likely to give the sort of feedback which encourages students to view failure as indicative of low ability, and therefore to become susceptible to LH. Some of the subjects in the present study were taught by a male. The role of teacher behaviour in determining self-worth motivation has yet to be examined.

ACKNOWLEDGMENTS—The author is grateful for the assistance of Iain Montgomery and John Davidson. Correspondence and requests for reprints should be sent to Dr. M. L. Craske, Department of Psychology, University of Tasmania, Hobart, Tasmania, Australia.

## REFERENCES

ARKIN, R. M., and MARUYAMA, C. M. (1979). Attribution, affect, and college exam performance. *J. educ. Psychol.*, 71, 85-93.

ABRAMSON, L. Y., SELIGMAN, E. P., and TEASDALE, J. D. (1978). Learned helplessness in humans: critique and reformulation. *J. abnorm. Psychol*, 87, 49-74.

ANDREWS, G. R., and DEBUS, R. L. (1978). Persistence and the causal perceptions of failure: modifying cognitive attributions *J. educ. Psychol.*, 70, 154-166.

BANDURA, A. (1981). Self-referent thought: a developmental analysis of self-efficacy. In FLAVELL, J. H., and ROSS, L. (eds.), *Social Cognitive Development: Frontiers and Possible Futures*. Cambridge: Cambridge University Press.

BEERY, R. (1975). Fear of failure in the student experience. *Personnel and Guidance J.*, 54, 190-203.

BREEN, L. J., VULCANO, B., and DYCK, D. (1979). Observational learning and sex roles in learned helplessness. *Psychol. Rep.*, 44, 135-144.

BROCKNER, J. (1979). The effects of self-esteem, success-failure, and self-consciousness on task performance. *J. Person. soc Psychol*, 37, 1732-1741.

BROWN, I., and INOUYE, D. K. (1978). Learned helplessness through modelling: the role of perceived similarity in competence. *J. Person. soc. Psychol.*, 36, 900-908.

BUTKOWSKY, I. S., and WILLOWS, D. M. (1980). Cognitive-motivational characteristics of children varying in reading ability· evidence of learned helplessness in poor readers. *J. educ. Psychol.*, 72, 408-422.

COOPERSMITH, S. (1981). *Self-Esteem Inventories*. Palo Alto, Calif.: Consulting Psychologists Press.

COVINGTON, M. V., and OMELICH, C. L. (1979). It's best to be able and virtuous too: student and teacher evaluative responses to successful effort. *J. educ. Psychol.*, 71, 688-700.

COVINGTON, M. V., and OMELICH, C. L. (1981). As failures mount: affective and cognitive consequences of ability demotion in the classroom. *J. educ. Psychol.*, 73, 798-808.

CRANDALL, V. C. (1969). Sex differences in expectancy of intellectual and academic reinforcement. In SMITH, C. P. (ed.), *Achievement-related Motives in Children*. New York: Russell Sage Foundation.

CRASKE, M. L. (1985). Improving persistence through observational learning and attribution retraining. *Br. J. educ. Pschol.*, 55, 138-147.

DE VELLIS, R. F., DE VELLIS, B. M., and MCCAULEY, C. (1978). Vicarious acquisitions of learned helplessness. *J. Person. soc. Psychol.*, 36, 894-899.

DIENER, C, and DWECK, C. S. (1978). An analysis of learned helplessness: continuous changes in performance, strategy, and achievement cognitions following failure. *J. Person. soc. Psychol.*, 36, 451-462.

DIENER, C., and DWECK, C. S. (1980). An analysis of learned helplessness: II The processing of success. *J. Person. soc. Psychol.*, 39, 940-952.

DWECK, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *J. Person. soc. Psychol.*, 31, 674-685.

DWECK, C. S., and BUSH, E. S. (1976). Sex differences in learned helplessness: I. Differential debilitation with peer and adult evaluators. *Dev. Psychol.*, 12, 147-156.

DWECK, C. S., DAVIDSON, W., NELSON, S., and ENNA, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom III. An experimental analysis. *Dev. Psychol.*, 14, 268-276.

DWECK, C. S., and GILLIARD, D. (1975). Expectancy statements as determinants of reactions to failure: sex differences in persistence and expectancy change. *J. Person. soc. Psychol*, 32, 1077-1084.

DWECK, C. S., and REPPUCCI, D. (1973). Learned helplessness and reinforcement responsibility in children. *J. Person. soc. Psychol.*, 25, 109-116.

FOWLER, J. W., and PETERSON, P. L. (1981). Increasing reading persistence and altering attributional style of learned helpless children. *J. educ. Psychol.*, 73, 251-260.

Frankel, A., and Snyder, M. L. (1978). Poor performance following unsolvable problems: learned helplessness or egotism? *J. Person. soc. Psychol.*, 36, 1415-1423.

Gatting-Stiller, I., Gerling, M., Stiller, K., Voss, B., and Wender, I. (1979). Anderungen der Kausalattribuierung und des Ausdaververhaltens bei misserfolgsmotivierten Kindern durch Modellernen. *Zeitschrift fur Entwicklungspsychologie und Padagogische Psychologie*, 4, 312-321.

Gody, D. S. (1978). Sex differences in reactance and learned helplessness. *Dissent. Abstr. Int.*, 38 (11-B), 5566.

Gollwitzer, P. M., Earle, W. B., and Stephan, W. G. (1982). Affect as a determinant of egotism: residual excitation and performance attributions. *J. Person. soc. Psychol.*, 43, 702-709.

House, W. C. (1980). Effects of knowledge that attributions will be observed by others. *J. Res. Person.*, 14, 528-545.

Ickes, W., and Layden, M. A. (1978). Attributional styles. In Harvey, J. H., Ickes, W., and Kidd, R. F. (eds.), *New Directions in Attribution Research.* Volume 2 New Jersey: Lawrence Erlbaum.

Keppel, G. (1982). *Design and Analysis: A Researcher's Handbook.* New Jersey: Prentice-Hall.

Kimball, O. M. (1973). Development of norms for the Coopersmith Self-Esteem Inventory: grades four through eight. Doctoral dissertation, Northern Illinois University, 1972. *Dissert. Abstr. Int.*, 34, 1131-1132.

Kokenes, B. (1974). Grade level differences in factors of self-esteem. *Dev. Psychol.*, 10, 954-958.

Kokenes, B. (1978). A factor analytic study of the Coopersmith Self-Esteem Inventory. *Adolescence*, 13, 149-155.

Le Unes, A. D., Nation, J. R., and Turley, N. M. (1980). Male-female performance in learned helplessness. *J. Psychol.*, 104, 255-258.

McNemar, Q. (1969). *Psychological Statistics.* New York: John Wiley.

Medway, F. J., and Venino, G. R. (1982). The effects of effort feedback and performance patterns on children's attribution and task persistence. *Contemp. educ. Psychol.*, 7, 26-34.

Nicholls, J. G. (1975). Causal attributions and other achievement-related cognitions: effects of task outcome, attainment value, and sex. *J. Person. soc. Psychol.*, 31, 379-389.

Nicholls, J. G. (1979). Development of perception of own attainment and causal attributions for success and failure in reading. *J. educ. Psychol.*, 31, 379-389.

Schunk, D. H. (1983). Ability versus effort attributional feedback: differential effects on self-efficacy and achievement. *J. educ. Psychol.*, 75, 848-856.

Schunk, D. H. (1984). Sequential attributional feedback and children's achievement behaviours. *J. educ. Psychol.*, 76, 1159-1169.

Schwarzer, R., and Jerusalem, M. (1982). Self-serving attributions after performance feedback. *Zeitschrift fur Entwicklungpsychologie und Padagogische Psychologie*, 14, 47-57.

Simon, W. E., and Simon, M. G. (1975). Self-esteem, intelligence and standardized academic achievement. *Psychology in the Schools*, 32, 97-100.

Spatz, K., and Johnston, I. (1973). Internal consistency of the Coopersmith Self-Esteem Inventory. *Educ. psychol. Measur.*, 33, 875-876.

Snyder, M. L., Smoller, B., Strenta, A., and Frankel, A. (1981). A comparison of egotism, negativity, and learned helplessness as explanations for poor performance after unsolvable problems. *J. Person. soc. Psychol.*, 40, 24-30.

Snyder, M. L., Stephan, W. G., and Rosenfield, D. (1978). Attributional egotism. In Harvey, J. H., Ickes, W., and Kidd, R. F. (eds.), *New Directions in Attribution Research.* Volume 2. New Jersey: Lawrence Erlbaum.

Weiner, B. (1972). *Theories of Motivation: From Mechanism to Cognition.* Chicago: Rand McNally.

Weiner, B. (ed.) (1974). *Achievement Motivation and Attribution Theory.* New Jersey: General Learning Press.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *J educ. Psychol.*, 71, 3-25.

Wilson, M., Seybert, J., and Craft, J. (1980). Human learned helplessness as a function of sex and degree of control over aversive events. *Bull. Psychonom. Soc.*, 16, 209-212.

Zimmerman, B. J., and Blotner, R. (1979). Effects of model persistence and success on children's problem-solving. *J. educ Psychol.*, 71, 508-513

Zimmerman, B. J., and Ringle, J. (1981). Effects of model persistence and statements of confidence on children's self-efficacy and problem-solving. *J. educ. Psychol.*, 73, 485-493.

(*Manuscript received* 10*th December*, 1986)

*Br. J. educ. Psychol.*, 58, 165-171, 1988

# PRE-SCHOOL CHILDREN'S PREFERENCES FOR DRAWINGS OF A SIMILAR COMPLEXITY TO THEIR OWN

By MICHAEL R. BROOKS, SHEILA M. GLENN AND W. RAY CROZIER
*(School of Psychology, Lancashire Polytechnic)*

SUMMARY. The drawings of people and houses by nursery school children were compared with their preferences for representations of various complexity of the same subjects. It was found that children significantly prefer representations which have most in common with their own drawings, thus indicating that children draw as they do through choice, and not because they have difficulty in producing more advanced representations.

## INTRODUCTION

THERE seems to be a significant degree of consensus among studies which have traced the development of children's drawings that (a) by the fourth year there is typically a transition from non-representational drawings which combine elementary forms and patterns to the attempt to depict people, houses etc., and that (b) these first representations characteristically contain many features which adults would consider to be errors. For example, drawings of people often lack distinct head and body, with arms and legs attached to one circular form, the so-called "tadpole figure" (Freeman, 1980).

There has been much less agreement about interpretation of these trends in development, and at present there are several rival theoretical positions. One class of explanations emphasises the child's intentions in making the drawing: Wilson and Wilson (1982) have argued that forms like the tadpole figure are a style of representation that has been influenced by the drawings of other children; Arnheim (1974) has proposed that children take forms from their available drawings repertoire like the circle to serve as equivalents for perceptually salient factors of the object; finally, a position that can be traced back to Luquet (1927) stresses the knowledge about properties of objects that outweighs perceptual information in choosing what to include in the drawing (Bremner, 1985). In opposition to these explanations Freeman (1980) has argued that children's depictions appear to have errors because of lack of cognitive and technical skills in the ability to produce drawings. For example, the tadpole figure is a function of memory limitations; the child produces the elements head, body, arms and legs in a fixed sequence but memory demands can result in elements in the middle of the sequence being forgotten.

Moore (1986) has recently investigated the question of whether depictions reflect internal representations of important features of objects or alternatively a lack of technical skills. She used a picture selection task and argued that if children prefer drawings of objects which are most similar to their own spontaneous drawings, then the implication is that children's drawings do reflect internal representations and are not constrained by production difficulties. Moore criticised previous studies in this area on a number of grounds including (a) whether the drawings used for ascertaining preference were similar in style to the children's own, (b) use of between-subject designs, (c) complexity of stimuli not well controlled, (d) extent of choice allowed. Controlling for these variables she found that in a sample aged from 4½ years to 9½ years children did prefer drawings of houses which had most features in common with their own drawings of houses.

The present study extends Moore's work in two main ways. Firstly, a younger age group is used; these children are likely to have more production problems. Secondly, drawings of people are also used. Representations of people are typically the earliest representational drawings children produce; moreover there is an extensive literature on the developmental sequence of people drawings.

Preparation of the drawings for the picture preference task was done partly by reference to previous literature (e.g., Gardner, 1980; Moore, 1986; Cox and Parkin, 1986), but also by inspecting spontaneous drawings of houses and people produced prior to the study by the children who were to participate. This was to ensure that the drawings used in preference tests were representative in style of this particular group, i.e., that complexity rather than style per se was the independent variable. For nursery school children aged between 3½ and 4½ years the following developmental patterns emerged: for people, first a circle is drawn with two lines at the bottom representing legs — the tadpole shape. Then eyes are added, followed by a mouth, nose, hair and arms. Fingers, feet and other features will then appear. Finally a body will be added. For houses the basic starting point is a square to which a door is first added followed by windows and a chimney. The roof appears relatively late.

Thus this study investigated whether, from the very first attempts at representational drawings of people and houses, children are satisfied with their own production, or whether they prefer more complex representations which they do not yet have the skill to produce.

## METHOD

### *Children*

The spontaneous drawings of people and houses of 24 nursery school children were studied to develop preference stimuli. Of these 17 children took part in the main study: nine girls and eight boys, mean age 4·0 years, standard deviation 3·5 months.

### *Design and materials*

Analysis of the children's spontaneous drawings prior to the experiment and information from the research literature generated three major distinguishing features for drawings of both people and houses. These were:

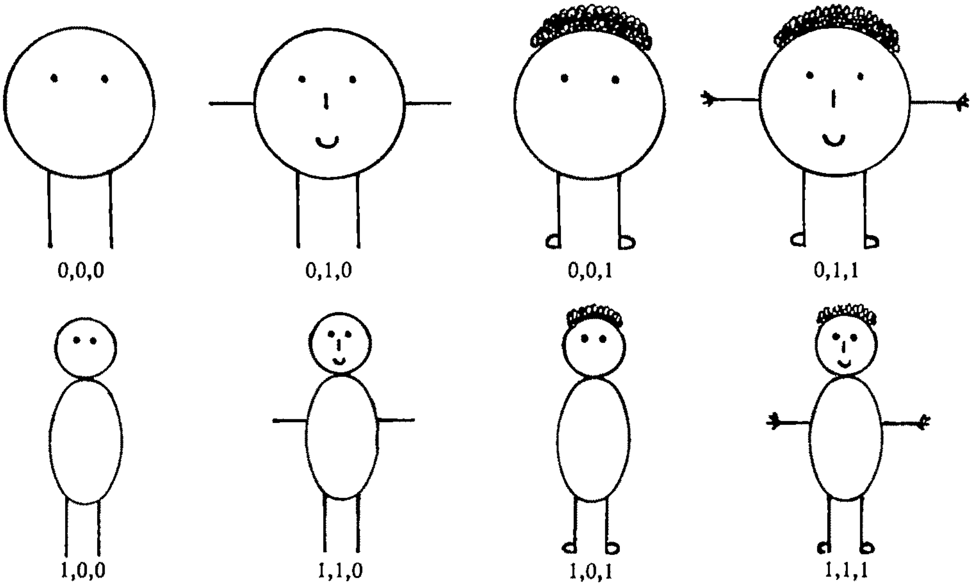
For people,
Feature 1 : body or tadpole
Feature 2 : nose, mouth, arms
Feature 3 : feet, fingers, hair

For houses,
Feature 1 : roof
Feature 2 : door
Feature 3 : windows, chimney

Standard representations of the stimuli were constructed in a style most typical of the children's drawing. Representations of every possible combination of the present and absent instances of the three features were drawn in black felt pen, each on a separate sheet of white A4 paper. This meant eight different versions of both the house and the person. The rectangular part of the house measured 3 × 3·5 ins. the triangular roof was 1·5 ins. in height. The tadpole people were drawn with a circle measuring 2·5 ins. in diameter. All types of people were 4 ins. in height and had the same facial expressions. The series of drawings used are shown in Figures 1 and 2.

FIGURE 1

SERIES OF 8 PEOPLE DRAWINGS USED IN PAIRED PREFERENCE TASK
Subscripts refer to presence (1) or absence (0) of features

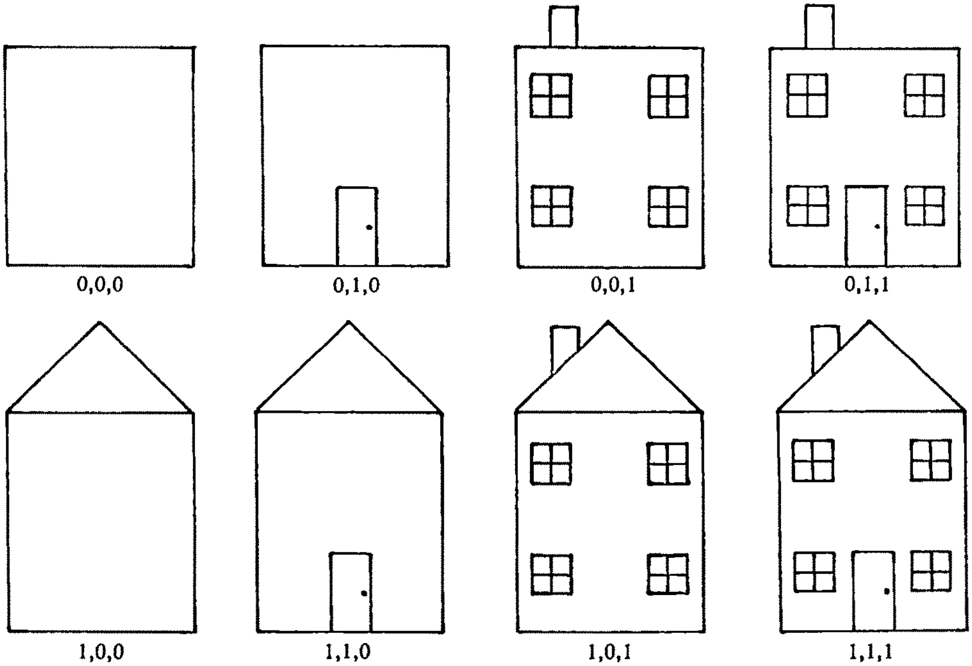


0,0,0 0,1,0 0,0,1 0,1,1

1,0,0 1,1,0 1,0,1 1,1,1

FIGURE 2

SERIES OF 8 HOUSE DRAWINGS USED IN PAIRED PREFERENCE TASK
Subscripts refer to presence (1) or absence (0) of features



0,0,0 0,1,0 0,0,1 0,1,1

1,0,0 1,1,0 1,0,1 1,1,1

A booklet was then prepared for each stimulus. Each booklet comprised all 28 possible comparisons of the eight people/house drawings (i.e., one booklet for people drawings, another for house drawings). The placement of each drawing on the left or right hand side of the page was balanced with the constraint that the same person/house did not appear on two consecutive pages. The drawings were rearranged and presented in different order for half the children to eliminate any position effects in the results. For the drawing task children were given a black felt tip pen and a blank sheet of white A4 paper.

### *Procedure*

The study was divided into two parts: a drawing task and a picture selection task. All children took part in both tasks. They were tested individually over a two-week period, with the drawing task, followed by the preference task on a later day. The experiment was conducted in a familiar separate corner of their nursery school. The children were used to the presence of the experimenter over a four-week period prior to the experimental sessions.

For the drawing task children were asked to draw a man/house on the paper provided. No help was given, but sometimes encouragement was necessary before some children started drawing.

For the preference task children were presented with a booklet which they were told contained "lots of drawings of men/houses done by some other children". The experimenter turned the pages and children were asked to choose one of each pair which was the best drawing of a man/house. For children who seemed unsure as to what to do, the differences between drawings in a pair were pointed out and children were asked which drawing they thought was best. Responses were recorded on a separate prepared sheet. The order of preference tests for people/houses was alternated for consecutive children.

## RESULTS

### *Drawings*

All but two of the children produced drawings with identifiable features. Two independent raters classified the drawings according to whether they did or did not possess the feature in question. There was 88 per cent agreement on people drawings, and 94 per cent on house drawings. Disagreements were settled by consensus. Table 1 shows the proportion of children drawing particular features.

TABLE 1

PROPORTION OF CHILDREN DRAWING PARTICULAR FEATURES

| | |
|---|---|
| *People* | |
| Feature 1 : Body | 0·24 |
| Feature 2 : Nose, Mouth, Arms (At least 2/3) | 0·82 |
| Feature 3 : Hair, Feet, Fingers (At least 2/3) | 0·59 |
| *Houses* | |
| Feature 1 : Roof | 0·12 |
| Feature 2 : Door | 0·76 |
| *Feature 3 : Windows or Chimney | 0·82 |

* Of the 10 children who drew either windows or chimney, 9 drew windows, and 1 a chimney.

### *Preferences*

To provide a measure of consistency of preferences, circular triads were calculated. A circular triad exists if A is preferred to B, and B to C, but C to A instead of A to C which would be consistent. Thus the higher the number of circular triads, the more inconsistent are subjects' judgments. Out of a highest possible score of 20, the mean number of circular triads was 10 for people drawings, and 9 for houses. This is comparable to Moore's (1986) score of 8 for 4½ year-olds. It should be noted that if children have no particular preference, (which could be the case if neither one of the pair is seen as a satisfactory representation) but are nevertheless asked to choose, then some degree of inconsistency is not surprising.

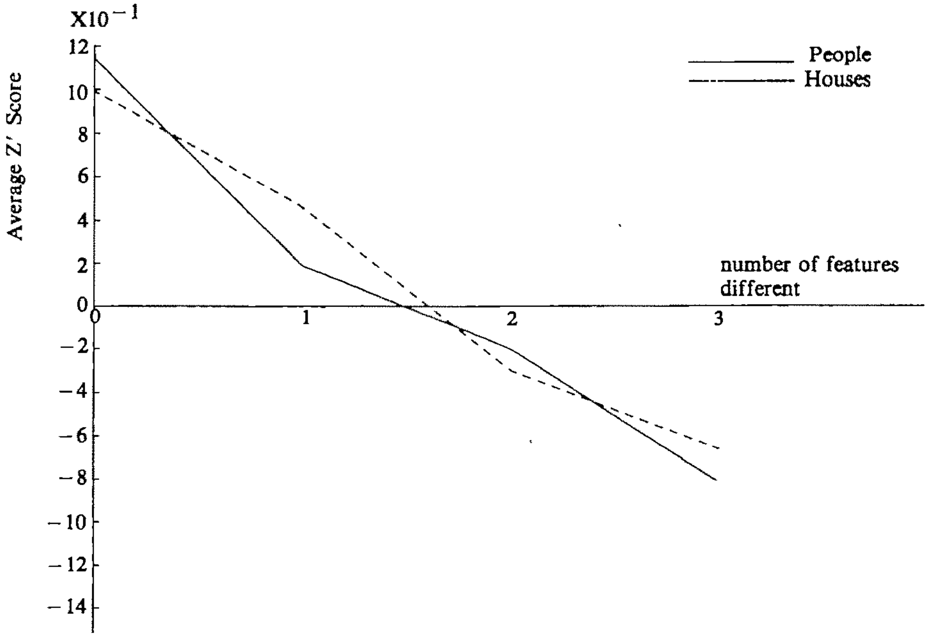
Choice scores were then obtained for each stimulus variation by counting the number of times each variation was judged "best". The choice scores were converted into Z scores, which were then converted into Z′ scores (Guilford, 1936). The Z′ scores are representative of the degree of preference by the children for each stimulus assuming a normal distribution of judgments.

For each child the mean preference score (Z′) was calculated for the stimulus drawings that differed from the child's own drawings by 0, 1, 2 and 3 features respectively. The mean Z′ score for the group for each of the four stages of difference in features was calculated for people and house stimuli, and is summarised in Figure 3.

It can be seen that children prefer drawings which are most similar to their own drawings of people and houses. It could be that preferences were obscured by the fact that drawings, which for example were one feature different from the child's own, could be both more or less complex, thus averaging results could be misleading. However, there was no evidence for this. Inspection of individual results

FIGURE 3

MEAN Z′ SCORES FOR DRAWINGS WHICH DIFFERED BY 0, 1, 2, 3 FEATURES FROM OWN DRAWINGS



showed that highest preferences were generally for drawings at a level similar to the child's own. In five instances (out of a possible 136) where this was not the case for people drawings, three involved preference for less complex drawings and two for more complex. Similarly for houses: of nine instances of higher preference four were for less complex and five for more complex. Thus there is no evidence to suggest systematic preference for drawings other than these similar to the child's own. Jonckheere's trend test (Jonckheere, 1954) showed that the ordering of preferences from 0 to 3 features different was significant for both people and houses ($P < 0{\cdot}001$).

For the people drawings two children produced drawings with no features, i.e., simply an enclosed shape. Their average $Z'$ scores for 0, 1, 2, 3 features difference were $-0{\cdot}2$, $0{\cdot}05$, $0{\cdot}2$ and $0{\cdot}05$ respectively. Similarly for house drawings three children produced unidentifiable drawings and their average $Z'$ scores were $-0{\cdot}43$, $-0{\cdot}4$, $0{\cdot}33$ and $0{\cdot}5$ with large between-subject variation. In other words, for children who were not yet producing identifiable drawings there was little evidence of systematic preference.

## DISCUSSION

The main finding from this study is very clear: children do prefer drawings of both people and houses which have most features in common with their own drawings. This suggests that children are satisfied with their own drawings, despite the presence of what from an adult viewpoint are errors. This confirms Moore's (1986) findings and suggests that even very young children, who are particularly likely to have production problems, prefer their own style of drawings. Furthermore, the result is confirmed for drawings of people, one of the earliest representations attempted by pre-school children.

The results are also consistent with a recent study by Cox and Parkin (1986). They found that reducing production demands on children (for example, by giving a copying task, a jig-saw task and a dictation task where children were instructed what to draw step by step) did not reduce their likelihood of drawing a tadpole representation of a person.

Why would children prefer their own style of drawing? In the light of the suggestion that the child's knowledge of an object is an important influential factor it could be that children simply try to convey what they decide to be the important features of the object (Moore, 1986). This suggestion fits in with the arguments proposed by other authors. Luquet (1927) argued that children selected and drew the details that were prominent in their knowledge of an object. Also Arnheim (1974) and Golomb (1981) have both suggested that it is in the nature of representational drawing not to copy reality but to invent equivalent forms instead. Hence, they argue the child will invent a form, which in their view stands adequately for the object. Moore herself found evidence that children who drew apparently insufficient representations of houses in fact intended the drawing to stand for the whole object. Moore also suggests, therefore, that development in drawing does not merely consist of the mastering of production abilities but also in the developmental change in the child's understanding and judgment of important properties of objects. Indeed, inspection of the children's drawings indicates that poor production ability (shown by a child who has difficulty in drawing complete circles and features) is not necessarily related to the production of simpler representations.

Of interest is the finding that the three children who produced unidentifiable drawings (scribblers) did not prefer the simplest representation and showed a great deal of between-subject variation in their preferences. This is consistent with the results of Taylor and Bacharach (1981) and they suggest an interactive process with the development of drawing being influenced by how children represent their world, but also influencing such representations. A longitudinal study would help to elucidate this relationship.

These results are of relevance to teachers of pre-school children, as whether children's behaviour is seen as "incorrect" or as giving insight into the way they think may have strong effects on teaching style. If children are satisfied with their own drawings and these reflect their representations of the world, then teachers' attempts to "correct" these may have potentially adverse effects on motivation and learning (e.g., Donaldson 1985). Of interest would be a study investigating teachers' beliefs about why children draw the way they do, and how this influences both teachers' interactions with children and children's subsequent drawings and enthusiasm for this activity.

ACKNOWLEDGMENTS.—Thanks are due to the staff and pupils of Stoneygate Nursery School, Preston.

Correspondence and requests for reprints should be addressed to Dr. S. Glenn, School of Psychology, Lancashire Polytechnic, Preston, PR1 2TQ, England.

## REFERENCES

ARNHEIM, R. (1974). *Art and Visual Perception.* Berkeley: University of California Press.

BREMNER, J. G. (1985). Provoked use of height in picture as a depth cue in young children's drawings. *Br. J. Dev. Psychol.,* 3, 95-98

COX, M. V., and PARKIN, C. E. (1986). Young children's human figure drawing: cross sectional and longitudinal studies. *Educ. Psychol.,* 6, 353-368.

DONALDSON, M. (1985). *Children's Minds.* London: Fontana.

FREEMAN, N. H. (1980). *Strategies of Representation in Young Children.* London: Academic Press.

GARDNER, H. (1980). *Artful Scribbles — The Significance of Children's Drawings.* London: Jill Norman Ltd.

GOLOMB, C. (1981). Representation and reality: the origins and determinants of young children's drawings. *Rev. Res. vis. Arts Educ.,* 14, 36-48.

GUILFORD, J. P. (1936). *Psychometric Methods.* New York: McGraw-Hill.

JONKHEERE, A. R. (1954). A test of significance for the relation between M rankings and K ranked categories. *Br. J. stat. Psychol.,* 7, 93-100.

LUQUET, G. H. (1927). *Le Dessin Enfantin.* Paris: Alcan.

MOORE, V. (1986) The relationship between children's drawings and preferences for alternative depictions of a familiar object. *J exp. Child Psychol.,* 42, 187-198.

TAYLOR, M., and BACHARACH, V. R. (1981). The development of drawing rules: metaknowledge about drawing influences performance on non-drawing tasks. *Child Dev.,* 52, 373-375.

WILSON, M., and WILSON, B. (1982). The case of the disappearing two-eyed profile: or how little children influence the drawings of little children. *Rev. Res. vis. Arts Educ.,* 15, 19-32.

(*Manuscript received* 17*th August,* 1987)

*Br. J. educ. Psychol.*, 58, 172-183, 1988

# PRESCHOOLERS' COGNITIONS OF FRACTIONAL UNITS

By ROBERT P. HUNTING
*(La Trobe University, Victoria, Australia)*

AND CHRISTOPHER F. SHARPLEY
*(Monash University, Victoria, Australia)*

SUMMARY. Problems requiring the sharing of discrete and continuous quantities, and problems involving the fractions ½, ⅓, and ¼ were individually presented to 206 preschool children. The children conceived of sharing in flexible and idiosyncratic fashion, showing common non-exhaustion of the material which was required to be shared and little evidence of anticipatory strategies prior to making subdivisions or checking and adjustment of resultant shares. While most children did not understand the fraction language used, there was evidence of a qualitative concept of one-half. A systematic procedure for sharing discrete objects was demonstrated by 60 per cent of the children. This procedure resulted in equal fractional units and could be a basis for meaningful fraction knowledge.

## INTRODUCTION

MUCH school mathematics is devoted to teaching concepts and procedures based on those units that form the core of whole number arithmetic, such as ones, tens, and hundreds. Other topics such as fractions and decimals demand new and extended understanding of units and their relationships. Researchers have noted how children's whole number ideas interfere with their efforts to learn fractions (Behr *et al.*, 1984; Streefland, 1984; Hunting, 1986). Hunting (1986) suggested that a reason why children seem to have difficulty learning stable and appropriate meanings for fractions is because instruction on fractions, if delayed too long, allows whole number knowledge to become the predominant scheme to which fraction language and symbolism is then related. There is some evidence which suggests that children can perform fraction-related tasks at ages prior to that at which these procedures are taught in school. Polkinghorne (1935) concluded from a study of 266 kindergarten, first, second, and third grade children that considerable knowledge of fractions is held prior to formal instruction in this topic, and Gunderson and Gunderson (1957) demonstrated that second graders had concepts and ideas about fractions that could be developed subsequently.

Both Kieren (1983) and Vergnaud (1983) have emphasised the primacy of partitioning for establishing fraction knowledge. Studies by Korbosky (1984), Miller (1984), Piaget *et al.* (1960), Pothier and Sawada (1983), and Smith (1985) contribute to our understanding of how children solve problems of partitioning continuous quantities when those quantities are regions. Piaget *et al.* (1960) studied children of ages 4 to 7, who were individually asked to cut a circular clay "cake" for different numbers of dolls. A number of conditions or characteristics of a unified fraction construct were proposed by those authors. Pothier and Sawada (1983) proposed a five-level theory of partitioning based on observations of children's attempts to subdivide continuous quantities represented by cakes of various dimensions and shapes. Detailed observations of the behaviour of 12 children aged 9-10 years as they attempted to partition linear, rectilinear and circular materials which varied across the three dimensions of unmarked, grid-marked, and adjustable material were reported by Korbosky (1984). Smith (1985) contrasted second and third graders' attempts to select correct partitions from a set of options with their attempts to generate solutions. Shapes used were circles, squares, and rectangles divided into three, five, six, and seven parts. Hiebert and Tonnessen (1978) found that children used different methods to subdivide discrete material such as penny candy compared to those used to subdivide continuous material. Miller (1984) studied young children's understandings of subdivision of sets, length, area, and volume. A study of children's partitioning behaviour in discrete quantity contexts (Hunting, 1981) showed that when singleton units were available, the nine 9- and 10-year-old children studied employed either anticipatory solutions based on whole number facts or conducted physical experiments involving various systematic and trial-and-error procedures for equalising shares. An operational scheme based on compensation logic was advanced to explain how children coordinate relations between the sizes and numbers of units in partitions where the units consisted of indefinite numbers of singletons.

Children aged 9 and 10 years have been found to possess (i) cognitive mechanisms which enable them to physically partition discrete quantities into equal shares, (ii) numerical processes for anticipating fractional unit sizes, and (iii) logic for relating different size units given the numbers of units to be compared, and (iv) halving and measuring processes for subdividing continuous material (Hunting, 1982, 1983b; Korbosky, 1984). However, little is known about the origins of such knowledge. It is possible that young children have sufficiently developed cognitive processes for dealing with problems involving subdivisions of quantities to allow fraction instruction to begin earlier. To determine if this were so, a careful study of younger children's responses to fraction tasks was designed to clarify what (if any) notions or skills in regard to fractions children hold, and what mental processes exist that may be used as bases for learning fractions.

Entry into the school system was considered to be an appropriate place to begin observing children. It is known that pre school children bring to school a wealth of informal and powerful strategies for dealing with mathematical problems (see, for example, Rea and Reys, 1970; Ginsburg, 1977; Brush, 1978; Gelman and Gallistel, 1978; Hendrickson, 1979). Teachers need to know the nature of this prior knowledge that children have so that instruction can be planned which is consistent with, and takes advantage of, such knowledge. The present study was designed to investigate the responses of preschool children who were presented with problems of partitioning both continuous and discrete quantities, and their conceptions of the fractions ½, ⅓, and ¼. Particular questions guiding the study were:

1. What kinds of behaviours do preschoolers display when solving partitioning problems in discrete and continuous quantity contexts?
2. What mental processes seem to govern such behaviours?
3. What do preschoolers understand about the fractions ½, ⅓, and ¼?
4. What cognitive structures exist, or have the potential to develop, that would facilitate fraction learning?

## METHOD

Two hundred and six children whose ages ranged from 3 years 4 months to 5 years 2 months (mean = 4·5; SD = 4 months) were interviewed individually during June 1985. Eighteen children were under 4 years of age, 96 children were in the age range 4·0-4·5, 91 were in the range 4·6-4·11, and one child was aged 5. The children were of predominantly middle class background and attended seven different kindergarten and preschool centres around inner city and outer suburban areas of metropolitan Melbourne.

An adaptation of the clinical method (Hunting, 1983a) was used, the features of which included (a) careful trialling and scripting of interview plans to anticipate as many contingencies as possible; (b) structuring of interview plans so that every opportunity was given for success on each problem, including problems of graded difficulty within most settings; (c) choice of problem settings that would arouse the child's interest and provide motivation for solving a problem; and (d) provision of appropriate assistance by the interviewer if the child appeared to encounter difficulty with a problem. A question would be repeated using alternative vocabulary if the interviewer felt the child was disadvantaged.

In a preparatory study the responses of 22 4-year-old children to a set of sharing and fraction problems were videotaped and analysed. Based on those data the interview problems were revised and modified to incorporate commonly observed responses. Interview plans were prepared which integrated a standard script including information and directions to the child, as well as response category checklists that would serve as data sources. All interviews were conducted by an experienced preschool teacher, who underwent a training programme supervised by the authors.

### *Problem descriptions and rationales*

Two sets of problems were prepared. The first set required children to share different types of materials between two, three, and four dolls. The second set of problems required children to represent the fraction ½, and if successful, the fractions ⅓ and ¼, in different contextual settings.

*Problem 1: Skipping Ropes.* A child was introduced to two dolls, each of height 125 mm, and told that "These dolls want to have a skipping rope each. Each rope has to be the same or the dolls will be unhappy". A piece of macrame string 300 mm long was placed in front of the child, who was then asked to show how the rope could be cut "so that it is all used up and each doll gets an even share". A small pair of scissors was available to the child. Children who experienced difficulty cutting were assisted by the interviewer, who either held the string steady or cut where the child pointed. After the rope had been cut and the pieces allocated, the interviewer asked if all the rope was used up and if the dolls would be happy with their ropes. A follow-up problem involving three dolls and a 450 mm length of macrame string was then given.

This problem was designed to investigate strategies used by young children for subdividing one continuous dimension (length). From the pilot work two main types of behaviours were observed as the children cut the pieces of string. These were (1) application of a single cut leading to exhaustive use of the material, and (2) two cuts, resulting in left-over material. Different procedures were observed in those children who made two cuts. Response categories were included in the interview plans based on these observations.

*Problem 2: Sharing the Crackers.* Twelve cracker biscuits were placed on the table, in a single pile, near three dolls, and the child told that "mother wants all the crackers to be shared up evenly because there will be a fuss if one doll gets more than another doll". After the crackers had been given out, another doll was introduced. The child was asked to "fix it so that each doll has a fair share".

Initial trialling with this problem suggested that distributing crackers to two dolls was easily within the capability of this age child, so in order to keep the interview as short as possible it was considered more instructive to observe how the child shared between three and four dolls.

*Problem 3: Sausages.* The interviewer rolled out a ball of modelling clay into a sausage approximately 200 mm long and 20 mm wide. The child was asked to "show me how to cut this sausage in half" with a knife which was provided. Before cutting the sausage, the child was asked how many pieces would be obtained, and how many cuts would be needed. If the appropriate response was given, other similar sausages were presented for cutting into quarters and thirds.

From the pilot study two main solution procedures were observed. One procedure was a halving strategy, which was often repeated. The other was to cut a series of small pieces beginning at one end of the sausage. This problem matched the "Skipping Ropes" problem because of the continuous material used.

*Problem 4: Swapcards.* A set of 12 picture swapcards was handed to the child, who was asked to "help me put half those cards in this envelope". After the child had completed the task, and the interviewer had noted the number of remaining cards, a follow-up question ("Have you put half the cards in the envelope?") was asked and checking behaviour noted. If the appropriate response was given, the problem was repeated for one quarter and one third, again using 12 swapcards. This problem was cast in a discrete quantity context, and hence matched the "Sharing the Crackers" problem.

The problems were conceived of as power instruments: not all children interviewed were expected to succeed on all aspects of each problem, and children were not given problems that were obviously beyond their capacity. The problems were administered to individual children in a quiet area of each kindergarten. Each interview averaged 15 minutes, with the pace of the interviews being controlled by the interviewer according to the child's responses. The interview plans made provision for the child's solution attempts to be recorded for later analysis.

## RESULTS

### *Problem 1: Skipping Ropes*

The children's attempts to produce skipping ropes for two dolls are presented first, followed by anticipatory and checking behaviours. Partitioning and checking behaviours are then presented for the three doll problem.

*Partitioning behaviour: two dolls.* Of interest was the number of cuts made by each child, subdivision accuracy in the case of one cut, types of strategies used in the two-cut case, and the nature of checking behaviour. Nearly all children (89 per cent) cut the string just once. Eleven children (5 per cent) made two cuts, and another 11 (5 per cent) made three or more cuts. For those who made a single cut a measure of the difference in length between the produced pieces showed that the ropes of 79 children differed by 15 mm or less, 117 children differed by 30 mm or less, and 166 children differed by 60 mm or less.

Strategies observed for the 11 children who made two cuts included four children who were observed to cut the rope into two unequal pieces and then subdivide the larger piece, one child who subdivided the smaller of the two unequal pieces, and six children who cut two pieces each less than 50 mm long and discarded the remainder. Three children were observed spontaneously to fold the string before determining where to cut.

After sharing the material each child was asked if all the rope had been used. Affirmative responses were given by 175 (85 per cent) children, somewhat fewer than the number of children who had produced two pieces of string with one cut.

The children were asked if they considered the dolls to be happy with their ropes, to which 185 (90 per cent) agreed. Checking behaviour was observed in 93 cases. Forty-seven children were observed to look again at the ropes they had allocated to the dolls, 41 placed their ropes together side by side to compare lengths, and five children placed their ropes end-on. Further action was taken by 14 children, who took up the scissors and adjusted their lengths of rope by trimming.

- *Partitioning behaviour: three dolls.* A surprisingly large number (77) of children made just one cut, 83 made two cuts, 32 made three cuts, and 14 children were observed to make four or more cuts. A common response observed in the case of a single cut was to give the shorter piece to one doll and indicate that the longer piece would be shared by the other two. Various procedures were observed in the behaviour of children who made more than three cuts, including strategies such as measuring the second and third pieces off against the first piece, adjustment by trimming as the pieces were being cut, spreading the rope across the three dolls resulting in similar pieces covering each doll with a left over piece at each end. Other children made many cuts and selected three pieces to give to the dolls. Pieces allocated were not always the most alike in length. Children in the two-cut case made less accurate subdivisions compared with the one-cut case, when differences between lengths are considered in relation to the total length of string. Comparisons between longest and shortest pieces produced showed length differences of 22·5 mm or less in seven children, 45 mm or less in 20 children, and 90 mm or less in 53 children.

The strategy preferred by most of the children who made three cuts was a progression of linear cuts beginning at one end of the material (27 of 32). Only three children used a repeated halving strategy.

*Anticipatory and checking behaviour: three dolls.* Four children were observed to fold the material prior to cutting. The question of whether the dolls would be happy with their skipping ropes was answered "yes" by 190, "no" by 12, with four children giving a nil or indeterminate response. Checking behaviour was observed in 56 children, of whom 36 checked visually, 15 arranged the ropes side by side, and five placed the ropes end-on. Seventeen children made adjustments to their ropes by trimming as a consequence of checking.

### *Problem 2: Sharing the Crackers*

Of interest in this problem were the spontaneous and prompted partitioning procedures used by the children to share 12 cracker biscuits between three, then four dolls, as well as the methods used to check the equality of the shares. Solutions to the three doll problem will be reported first.

*Partitioning behaviour: three dolls.* Although the children were asked "Can you share up all the crackers so each doll has the same?" only 56 (27 per cent) distributed all 12 crackers spontaneously. Many children (111) stopped after allocating three crackers. Sixteen (8 per cent) children stopped after giving out six crackers, and the same number after giving out nine crackers. To prompt those who stopped short of distributing the set quantity the interviewer asked: "Are all the crackers given out to the dolls?". These children were encouraged to complete the task.

Systematic and unsystematic procedures were used to allocate crackers to dolls. Ninety (44 per cent) of the children used a systematic procedure in which a single cracker was given to each doll in turn. After each doll had received a cracker the process began again in a cyclic pattern until all crackers were apportioned equally. A more efficient variation of this procedure was observed in the behaviour of another 33 children (16 per cent), who adopted a many-to-one allocation of crackers to dolls for at least one cycle. The most common many-to-one methods were an initial allocation of lots of three followed by an allocation of a single cracker (11 children), and an initial allocation of lots of two followed two cycles of single crackers (eight children). Children observed to use non-systematic procedures would allot crackers to dolls either one-to-one, but not in rotation, or place crackers sometimes two at a time, sometimes one at a time. There were 20 children (10 per cent) who created equal shares as a result of non-systematic methods. However the majority of children (N = 55; 27 per cent) who used non-systematic procedures produced unequal shares.

*Checking behaviour: three dolls.* After the crackers were given out the children were asked if each doll got an even share, to which 179 children responded yes, 19 responded no, and seven made a nil or indeterminate response to the question. A follow-up question: "How can you tell?" attempted to uncover the kinds of checking methods available to the children (see Table 1). Most common methods observed were numerical justifications without overt counting, point counting of piles of crackers, and visual comparisons involving eye movements only. Two children slid piles of crackers together for direct comparisons, and 19 children gave tautological responses saying that the piles were even.

TABLE 1
SHARING CRACKERS BETWEEN THREE DOLLS

| | | CHECKING BEHAVIOUR | | |
|---|---|---|---|---|
| | Because they're even | Because they're uneven | Point counts crackers | Slides piles together |
| Frequency (N=206) | 19 | 1 | 26 | 2 |
| (%) | (9) | (0) | (13) | (1) |
| | Visual comparison | Numerical justification without overt counting | | Other |
| | 30 | 32 | 96 | |
| | (15) | (16) | (47) | |

*Partitioning behaviour: four dolls.* Regardless of whether each of the three dolls received the same number of crackers, the next problem was to re-apportion the 12 crackers between four dolls. Successful methods were observed in 94 children. Of these the most common can best be described as trial and error whereby, for example, two crackers would be taken from one doll and given to the fourth, then various adjustments made between piles until each pile was equalised. One cracker was taken from each existing pile by 22 children. Eleven children restored all the crackers back to a single pile before proceeding. Seven of these children then used a systematic one-to-one procedure; the other four used experimental methods (see Table 2). Unsuccessful attempts were made by 112 children. The most common response observed (32 children) was movement of two crackers from one of the existing dolls to the fourth doll. Twenty-seven of these children took two crackers from the doll closest to the newcomer. Another common response was removal of two crackers uniformly from each existing doll. For this strategy 16 children gave the new doll two crackers and put aside the remaining four, and three children left the new doll with six crackers. Other responses included removing two crackers from each of two existing dolls and giving three to the new doll, removing one cracker only for the new doll, and placing crackers around or adjacent to the dolls in various arrangements and patterns.

TABLE 2
SHARING CRACKERS THREE DOLLS TO FOUR DOLLS

| DISTRIBUTION PROCEDURES | Frequency (N=206) | (%) |
|---|---|---|
| Restores crackers to single pile and reallocates systematically | 7 | (3) |
| Restores crackers to single pile and reallocates unsystematically | 4 | (2) |
| Redistributes by trial and error | 61 | (30) |
| Takes one cracker from each doll | 22 | (11) |
| Other | 112 | (54) |

*Checking behaviour: four dolls.* In response to the question: "Has each doll got an even share?" 186 children said yes, 15 said no, and four gave nil or indeterminate responses. When asked to justify their decision, 44 children gave a numerical reason in which there was no overt counting activity, 15 point-counted the piles, 26 children were observed to make a visual check using eye contact only, and three physically moved piles together in order to make direct comparisons. A large number of children (118) either did not respond, said they didn't know, or offered reasons like "because they've all got crackers".

### *Problem 3: Sausages*

In this problem children were asked to cut a modelling clay sausage in half, and if successful, to cut similar sausages into quarters and thirds. Of interest in this problem were (1) responses of the children to preliminary questions designed to reveal the extent of anticipatory thought or prior knowledge of strategies for solving the given problem, and (2) partitioning behaviour as the children solved the problem.

*Anticipatory behaviour: one-half.* Before allowing the child to commence cutting, two questions were asked. These were: "How many pieces will we get?" and "How many cuts will you make?". Table 3 shows that 64 children predicted they would obtain two pieces. Other popular predictions were four pieces (41 children), and three pieces (29 children). For the question of how many cuts there was a variety of opinion with four cuts the most popular response (51 children), followed by three cuts (39 children), two cuts (22 children), and five cuts (19 children).

TABLE 3
HALF A SAUSAGE

PREDICTING PIECES

| Number of pieces | 1 | 2 | 3 | 4 | 5 | 6 | 7 or more | nil response |
|---|---|---|---|---|---|---|---|---|
| Frequency (N=206) | 10 | 64 | 29 | 41 | 18 | 8 | 4 | 32 |
| (%) | (5) | (31) | (14) | (20) | (9) | (4) | (2) | (16) |

PREDICTING CUTS

| Number of cuts | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 or more | nil response |
|---|---|---|---|---|---|---|---|---|---|---|
| Frequency (N=206) | 23 | 22 | 39 | 51 | 19 | 7 | 2 | 3 | 5 | 35 |
| (%) | (11) | (11) | (19) | (25) | (9) | (3) | (1) | (1) | (2) | (17) |

*Partitioning behaviour: one-half.* Appropriate responses were observed in the case of 73 children, who cut the sausage once near the mid-point and then stopped. Children who began at one end of the sausage and executed a succession of cuts numbered 118. Most commonly observed number of cuts were four (23), three (21), five (20), and seven (17). Eight children cut the sausage at about the mid-point, and continued to "halve" each segment so created.

*Quarters and thirds.* The 73 children who succeeded with the previous task were asked: "Can you cut this (another) sausage up into quarters for me?" Before cutting they were asked to say how many pieces and how many cuts they expected to make. Twenty children did not respond to the question of how many pieces. Three and four pieces were the most popular predictions (17 children each); nine children said five pieces (see Table 4). The question of how many cuts was not answered by 22 children. Sixteen children responded with three cuts, 11 children said four and five cuts respectively. The most popular strategy used was a succession of cuts beginning at one end. Of the 56 children who did this 15 made four cuts, 10 made three, eight made two, and eight made five cuts. Nine children cut once near the mid-point and then stopped.

TABLE 4
QUARTER OF A SAUSAGE

PREDICTING PIECES

| Number of pieces | 1 | 2 | 3 | 4 | 5 | 6 | 8 | nil response |
|---|---|---|---|---|---|---|---|---|
| Frequency (N = 75) | 2 | 6 | 17 | 17 | 9 | 1 | 3 | 20 |
| (%) | (3) | (8) | (23) | (23) | (12) | (1) | (4) | (27) |

PREDICTING CUTS

| Number of cuts | 1 | 2 | 3 | 4 | 5 | 6 | 8 or more | nil response |
|---|---|---|---|---|---|---|---|---|
| Frequency (N = 75) | 6 | 3 | 16 | 11 | 11 | 2 | 4 | 22 |
| (%) | (8) | (4) | (22) | (15) | (15) | (3) | (5) | 30 |

PARTITIONING BEHAVIOUR

| | Linear series of cuts | Cuts near mid-point and stops | Continuation of halving | Other |
|---|---|---|---|---|
| Frequency (N = 73) | 56 | 9 | 2 | 6 |
| (%) | (77) | (12) | (3) | (8) |

(Total % exceeds 100 because of rounding)

The 10 children who produced four pieces were given another similar sausage and asked to cut it up into thirds. No child predicted three pieces and only two children predicted two cuts. Eight children were observed to make a series of cuts beginning at one end. No child was successful.

### *Problem 4: Swapcards*

A set of 12 picture swapcards was handed to the child, who was told: "Here are some swapcards. We need to save half of them for another time. Can you help me put half those cards in this envelope?". After completing the task the child was asked: "Have you put half the cards in the envelope?". Of interest in this problem was the number of cards placed in the envelope, and the presence and nature of checking behaviour. The problem was repeated using one-quarter for children who succeeded.

*Partitioning behaviour: one-half.* A substantial number of children (73) placed all 12 cards in the envelope provided. Table 5 shows that while only 23 children responded appropriately, many children's responses cluster around six suggesting a qualitative or approximate notion of half.

*Checking behaviour: one-half.* There was agreement by 185 (90 per cent) children that half of the swapcards had been placed in the envelope. Checking behaviour was noted in 40 cases only. Only in two cases, however, did the interviewer's prompt lead to appropriate adjustments. In these cases the children counted the number of cards in each quantity. For the children who had placed all 12 cards in the envelope the most common response was just to look at the envelope's contents. Children who had divided their cards into two unequal sets either just looked at their cards, looked and spread the cards out, or added or subtracted cards from the envelope without over counting.

TABLE 5
Half of 12 Swapcards

| | CARDS PLACED IN ENVELOPE | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of cards | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| Frequency (N = 206) | 73 | 1 | 4 | 4 | 20 | 18 | 23 | 20 | 20 | 13 | 6 | 4 |
| (%) | (35) | (0) | (2) | (2) | (10) | (9) | (11) | (10) | (10) | (6) | (3) | (2) |

*One-quarter.* Of the 23 children who succeeded in placing six cards in the envelope, none responded appropriately for the problem involving a quarter, although 22 children agreed that they could solve the problem.

## DISCUSSION

The first two problems focused on sharing quantities between dolls. In the skipping ropes problem continuous material was used, while discrete material was used for the crackers problem. For many children sharing meant that each recipient received a portion of the material to be shared, without the necessity to distribute all the available material. The sum of the parts did not need to equal the original whole. Non-exhaustion of material to be shared was observed in both the skipping ropes and crackers problems. In the skipping ropes problem it was expected that the data might support a correspondence between an action of subdivision and a recipient — leading to the same number of cuts as dolls. Piaget *et al.* (1960) had noted confusion between number of cuts and number of pieces produced. Neither the two-doll nor the three-doll case provided strong evidence for this. An exception to the phenomenon of non-exhaustive distribution was the case of sharing rope between two dolls. Here most children made just one cut, and in so doing demonstrated considerable accuracy. While the data seem to indicate that, for sharing rope between three dolls, accuracy dropped for two cuts, it is possible that some children who might have stopped at two cuts, decided to continue on cutting in order to produce pieces less variable in length. If this were the case the problems as designed may have placed children in conflict between wanting to use all the material and wanting to produce equal parts. Some children may not have been able to remember both these conditions. The crackers problems would be less problematic in this regard, because the numbers of crackers allotted could be adjusted. Non-exhaustive distribution in the context of problems of sharing quantities is not surprising when one considers possible situations familiar to children where it is quite usual for recipients to be given equal shares without apportioning all the available quantity.

(For example, in the home it is common for a dessert to be eaten over two meals.) It is possible that children whose distributions were non-exhaustive attempted to keep the amounts allocated proportional to the sizes of the dolls.

In the second set of problems where the children were asked to represent half of a continuous quantity (sausages) and half of a discrete quantity (swapcards), there is some evidence to support the view that while many children did not show precision in their solutions, a qualitative conception of half existed which would precede more refined quantitative knowledge. Strongest support is found in the children's responses to the swapcards problem, where a single partition was created whose parts were unequal. The large number of children producing a linear series of cuts in the sausages problems may have occurred because the act of cutting, as a pleasurable experience, became an end in itself. Also the word "cut" may have been the stimulus that precipitated action rather than the word "half". A possible development of meaning for the word half, consistent with these data, would seem to be (a) half as an act of subdivision, including multiple subdivisions, (b) half as a single subdivision, resulting in two distinct quantities, (c) half as a single subdivision directed toward equality of quantities produced, and (d) half as one of two identical elements in thought, each element possibly a collection of units, and representable with physical material. Overall, few children interviewed evidenced mature knowledge of one-half. Little understanding of quarters and thirds was found.

A feature of the children's behaviour as they attempted to solve the problems of sharing and the problems involving fractions was a lack of anticipatory strategies and a lack of spontaneous verification or checking strategies. Anticipation of the likely outcome of an event or problem is a quality of intellectual activity important for successful, efficient problem solving. It may involve recalling an appropriate solution strategy remembered from previous experience, or performing operations in advance of actions that will serve to guide those actions to a successful conclusion. For the fractions problems lack of understanding of the language used may have thwarted such behaviour. In the sharing problems, where the given task was generally well-understood, most children proceeded to initiate a cut or subdivision with little or no preliminary reflection on the requirements of the task. Children's approaches to the skipping ropes problem exemplify the point. Almost in one motion a child would take the scissors in one hand, hold the piece of rope in the other, and directly begin to cut the rope. Very few children were observed to visually estimate or adjust the lengths of segment that resulted from placement of the scissors at a particular point, and other measuring strategies such as folding were rarely used. Similarly, after the portions had been given out, few children made adjustments after the interviewer asked if the dolls would be happy.

Explanations for the lack of anticipatory behaviour and checking strategies may be that the problems presented were quite dissimilar to other experiences previously encountered, so that they were not able to draw on prior knowledge; the children did not consider sharing to demand equal quantities; or the problems as communicated did not clearly emphasise the necessity for equal quantities to be determined. It is also possible that the children's notions of equality have not developed with respect to the need to produce equal quantities. A different task in which the children were recipients may have led to increased concern for equality.

Problematic situations involving sharing were considered to be rich contexts for laying bare the thinking of children as they partitioned quantities into fractional units. What has emerged from this study is that young children think about sharing in flexible and idiosyncratic fashion. For many children sharing does not necessarily mean "use it all up", nor does it necessarily mean that each recipient will be allocated a portion. That three dolls can share two skipping ropes implies that sharing has a quality of intermittent participation or interaction with the commodity spaced over a time period rather than simultaneous ownership requiring subdivision. One child dealt ingeniously with such an outcome by suggesting that the dolls swap the pieces around, possibly reflecting an adult suggestion previously experienced. The children studied did not universally believe that the quantity as given is absolute. If the situation demands it, it is reasonable to expect more material to be provided from somewhere, or, alternatively, to ignore some of what has been provided. Amounts that are allocated may depend on what the child considers appropriate according to factors other than exhaustion of the whole or creation of equivalent sub-units. For example, one child cut fairly small pieces of rope for the dolls because he considered that the size of the dolls warranted only small pieces.

Earlier research (Hunting, 1983b), which suggested that children's poor understanding of basic fraction concepts is not due to a lack of appropriate cognitive operational schemes for representing, relating, and transforming fractional units in concrete settings, is supported by these data. While few of these preschoolers seemed to understand the concept of one-half (much less one-third or one-fourth), 123 (60 per cent) of them demonstrated a powerful systematic algorithm in the crackers problem leading to the creation of accurate equal fractional units. The method involved three nested actions. The primary action was a matching of cracker to doll, and was repeated until each doll had a cracker. A cycle of matching cracker to doll for each doll present was the secondary action. After the cycle had been completed and there were crackers remaining to be shared, then the cycle was repeated. The third element of the algorithm was the iteration of the cycle. As has been noted in previous research (Hunting, 1981), while the child who uses this systematic method does not necessarily know at the end of it how many crackers each doll has received, the method guarantees that each doll receives an equal number of crackers. Such an algorithm would be an ideal action-meaning base for the mathematical language and symbolism used to represent numbers we know as fractions, particularly unit fractions such as ½, ⅓, and ¼. Sophisticated counting knowledge was not required to produce the required units, although counting could be used to verify solutions. Children who do not possess procedures for subdividing physical material may still have important notions about fractions, from observations of outcomes of others' sharing actions. Further research is needed to investigate the relation between knowledge based on personal actions of partitioning, and knowledge based on judgments of the actions of others (as for example, Davis and Pitkethly (1987), who have studied second grade children's views of pre schoolers' sharing behaviour using edited video segments).

This study shows that potential for learning fractions is present in children at an early age. Specifically, instructional activities which encourage exhaustive distributions, together with discussions about the relative merits of systematic versus non-systematic sharing procedures for making equal units, would be valuable in preparing children for further learning. As children become more proficient in allocating shares, fraction vocabulary such as one-half, one-third, one-fourth (or one quarter), halves, thirds, and fourths (or quarters) could be associated with the appropriate actions.

Correspondence and requests for reprints should be addressed to R. P. Hunting, School of Education, Centre for the Study of Curriculum and Teacher Education, La Trobe University, Bundoora, Victoria, Australia 3083.

## REFERENCES

BEHR, M. J., WACHSMUTH, I., POST, T., and LESH, R. (1984). Order and equivalence of rational numbers: a clinical teaching experiment. *J. Res. Math. Educ.*, 15, 323-341.

BRUSH, L. R. (1978). Preschool children's knowledge of addition and subtraction. *J. Res. Math. Educ.*, 9, 44-54.

DAVIS, G., and PITKETHLY, A. (1987). Cognitive aspects of sharing. Paper presented at the Workshop on Metacognition, Australian Council for Educational Research, Hawthorn, Victoria, October 24.

GELMAN, R., and GALLISTEL, C. R. (1978). *The Child's Understanding of Number*. Cambridge: Harvard University Press.

GINSBURG, H. P. (1977). *Children's Arithmetic: The Learning Process*. New York: D. Van Nostrand Co.

GUNDERSON, A. G., and GUNDERSON, E. (1957). Fraction concepts held by young children. *Arith. Teacher*, 4 (4), 168-174.

HENDRICKSON, A. D. (1979). An inventory of mathematical thinking done by incoming first-grade children. *J. Res. Math. Educ.*, 10, 7-23.

HIEBERT, J., and TONNESSEN, L. H. (1978). Development of the fraction concept in two physical concepts: an exploratory investigation. *J. Res. Math. Educ.*, 9, 374-378.

HUNTING, R. P. (1981). The role of discrete quantity partition knowledge in the child's construction of fractional number. (Doctoral dissertation, University of Georgia, 1980.) *Diss. Abstr. Int.*, 41, 4308A-4390A. (University Microfilms No. 8107919.)

HUNTING, R. P. (1982). Qualitative compensation thought in children's solutions to fraction comparison problems. In IRONS, C. J. (Ed.), *Research in Mathematics Education in Australia 1981* (Vol. 2). Kelvin Grove: Brisbane College of Advanced Education, Mathematics Education Research Group of Australasia.

HUNTING, R. P. (1983a). Emerging methodologies for understanding internal processes governing children's mathematical behaviour. *Aust. J. Educ.*, 27, 45-61.

HUNTING, R. P. (1983b). Alan: a case study of knowledge of units and performance with fractions. *J. Res. Math. Educ.*, 14, 182-197.

HUNTING, R. P. (1986). Rachel's schemes for constructing fraction knowledge. *Educ. Stud. Math.*, 17, 49-66.

KIEREN, T. E. (1983). Partitioning, equivalence, and the construction of rational number ideas. In ZWENG, M. (Ed.), *Proceedings of the 4th International Congress on Mathematical Education*. Boston: Birkhauser.

KORBOSKY, R. K. (1984). Partitioning continuous quantities. *Collected Papers of the National Conference of the Australian Association for Research in Education*. Perth, WA: Research Branch, Education Department of W.A.

MILLER, K. (1984). Child as the measurer of all things: measurement procedures and the development of quantitative concepts. In SOPHIAN, C. (Ed.), *Origins of Cognitive Skills*. Hillsdale, N.J.: Erlbaum.

PIAGET, J., INHELDER, B., and SZEMINSKA, A. (1960). *The Child's Conception of Geometry*. New York: Basic Books.

POLKINGHORNE, A. R. (1935). Young children and fractions. *Childhood Educ.*, 11, 354-359.

POTHIER, Y., and SAWADA, D. (1983). Partitioning: the emergence of rational number ideas in young children. *J. Res. Math. Educ.*, 14, 307-317.

REA, R. E., and REYS, R. E. (1970). Mathematical competencies of entering kindergartens. *Arith. Teacher*, 17, 65-74.

SMITH, J. (1985). Children's conceptual abilities in the partitioning task. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April.

STREEFLAND, L. (1984). Unmasking N-Distractors as a source of failures in learning fractions. In SOUTHWELL, B. (Ed.), *Proceedings of the Eighth International Conference for the Psychology of Mathematics Education*. Sydney: Mathematical Association of New South Wales.

VERNGAUD, G. (1983). Multiplicative structures. In LESH, R., and LANDAU, M. (Eds.), *Acquisition of Mathematics Concepts and Processes*. New York: Academic Press.

(*Manuscript received 1st June*, 1987)

*Br. J. educ. Psychol.*, 58, 184-190, 1988

# THE LONG TERM EFFECTIVENESS OF PARENTAL INVOLVEMENT IN READING: A FOLLOW-UP TO THE HARINGEY READING PROJECT

By JENNY HEWISON
*(Department of Psychology, University of Leeds)*

SUMMARY. The Haringey Reading Project was a two-year educational intervention, which sought, firstly, to increase the amount of parental help given to 6-8 year-old children learning to read, and secondly, to evaluate the impact of that help on the children's reading performance. As reported previously, at the end of the project children who had taken part in the parent involvement exercise were reading better than comparable children who had acted as controls. No such improvement was apparent in children given extra reading help, on a small group basis, from a supplementary teacher in school.

Three years after the end of the project, the children were tested in reading by the LEA, as part of their routine monitoring and screening exercise for children transferring from junior to secondary schools. Analysis of these data revealed that the beneficial effects on reading performance of having taken part in the parent involvement exercise were still clearly apparent three years later. Again, no clear pattern of benefit could be seen in the children who had received extra teacher help as part of the project.

## INTRODUCTION

IN 1980, Hewison and Tizard published a study of social background influences on reading failure, which identified active parental help as a major correlate of reading success, and suggested that the relationship might be causal.

In an attempt to investigate this relationship further, the Haringey Reading Project sought to increase the amount of reading help given to 6-8 year-old children by their parents (specifically, by asking the parents to listen to their children read), and to evaluate this intervention in terms of its impact, if any, on the reading progress of the children (Tizard, *et al.*, 1982). At the end of the two-year intervention period, the reading attainment of the project children (measured using the NFER "A") was significantly higher than that of controls. At a follow-up one year later, (using the NFER "BD"), project children in one of the two participating schools had maintained a statistically significant lead over controls; in the other school, this lead had been reduced and no longer reached statistical significance. Funding having expired, no further follow-up testing was carried out on behalf of the project.

The existence of any long-term benefits from special reading projects, has, in the past, proved difficult to establish (Hewison, 1982; Gittelman, 1985). It is therefore very fortunate that an opportunity has arisen to examine the longer-term progress of children who took part in the Haringey project, using reading data collected and generously made available by the Haringey Education Authority. This paper reports on the reading performance of children from the Haringey Project, tested as part of the borough's routine monitoring and screening exercise, three years after the end of the intervention period (i.e., two years after the follow-up mentioned above), when the children were about 11 years old.

The original Haringey Reading Project was mounted in six schools. Two of these housed the parent involvement project, two housed an intervention which consisted of giving supplementary reading instruction from a qualified teacher, and two provided entirely control information. Substantial between-schools differences reduced the usefulness of the last mentioned control data, and the main analyses used to evaluate the project were based on within-school comparisons. For the present very modest exercise, it was decided to follow this pattern again and to use the limited resources available to hunt down data for as many children as possible from the first four schools, at the expense of collecting information from the last two.

## METHOD

The London Reading Test was given to all final year juniors in Haringey in the summer before their transfer to secondary school. The group test was administered and scored by school staff, following the procedures specified in the Test Manual (ILEA, 1980). No further details are available about the precise arrangements for testing; some staff may have known which children had taken part in the intervention and which had not, but after three years and many staff changes, together with the usual re-assortment of children and teachers across classes, it is unlikely that strong systematic biases would have been introduced from this source.

The London Reading Test was designed to measure reading comprehension. In a recent test review, Bennett (1984) noted that:

> "Comprehension of the first and second passages is tested by the cloze technique. . . . The third passage is designed to stretch the more able children and here comprehension is tested by a series of questions covering the whole range of comprehension skills following Barrett's Taxonomy of Reading Comprehension" (p. 215).

The same reviewer commented that scoring procedures for the test were "clear and extensive" (p. 217); so again it seems unlikely that major distortions would have been introduced into the Haringey data by scorers' knowledge of project group membership.

Two standardisation tables are provided for the London Reading Test, one based on a sample of London children, and the other on a nationally representative sample. Since the overall reading standard of the former group was lower than that of the latter, a particular raw score translates into a higher standardised score if the London norms are used than if the national norms are used. In the Haringey exercise, school staff were instructed to tabulate both standardised scores, the "London" and the "National", beside each child's raw score.

The author was given access to the original data sheets submitted to the Education Office by schools. Her task was to extract the reading data for as many children as possible who had taken part in the intervention project, either in the experimental or control groups.

For children who were still attending the same school this was relatively straightforward. Children who had spent the intervention period in one school, but had subsequently moved on, were traced as far as was practicable. This amounted to scrutinising the lists of schools in the same part of the Borough, since it was known from experience that very local moves were common. Time and other practical constraints did not permit an exhaustive search of the lists of all borough schools, and, undoubtedly, a small number of available children will have been missed by the procedure adopted.

## RESULTS

At the end of the intervention period, data were available on 137 children from the two parent involvement schools, and 111 from the two "extra teacher help" schools. Data for 110 of the former group (80·3 per cent) and 90 of the latter (81·1 per cent) were traced and used in the three-year follow-up exercise. (Not all of these children were included in the one-year follow-up previously reported, since that was based on reading testing carried out for the project, and hence limited to the original six schools.) Data for 11 children were obtained from schools other than the one attended during the project. The data for these children, three who had taken part in the parent involvement exercise, one who had received supplementary teaching help and seven controls are included in the analyses reported below.

### *(a) Comparison of mean scores obtained at the three-year follow-up by children from the project's experimental and control groups*

Before the two-year intervention project began, within-school comparisons of children allocated to experimental and control groups had shown no significant group differences in reading performance.

TABLE 1

MEAN SCORES ON THE LONDON READING TEST, THREE YEARS AFTER THE END OF THE INTERVENTION

| | | Standardised score on London Reading Test (national norms) | | |
|---|---|---|---|---|
| School | Group | Mean | SD | N |
| 1 | Parent involvement | 99·8 | 8·37 | 18 |
| | Control | 92·15 | 11·85 | 40 |
| 2 | Parent involvement | 102·0 | 14·68 | 23 |
| | Control | 97·7 | 15·68 | 29 |
| 3 | Extra teacher help | 98·8 | 15.30 | 14 |
| | Control | 98·4 | 11.70 | 38 |
| 4 | Extra teacher help | 98·9 | 13·57 | 20 |
| | Control | 95·0 | 10·58 | 18 |

Table 1 summarises the performance of the various groups on the London Reading Test (national norms), administered three years after the end of the intervention period, when the children were about 11 years old.

A two-way analysis of variance (School x Project Group) carried out on the data from the two schools which had housed the parent involvement project yielded one significant main effect, that of Project Group (F = 5·16; df = 1,106; P = 0·025). Neither the interaction term, nor the main effect of school, reached statistical significance.

A comparable analysis carried out on the data from the two schools in which some children had received extra teacher help yielded no statistically significant results.

### *(b) The relationship between initial (pre-project) reading score and performance at follow-up*

Following the procedure adopted in the first evaluation study, the Southgate Reading Test scores obtained when the children were 5-6 years of age, and in the middle year of the infants school, were used to divide the sample into three initial reading bands: scores of 12 or below, scores of 13-19, and scores of 20-30.

A three-way analysis of variance (School x Project Group x Initial Reading Band) carried out on the London Reading Test scores from the two parent involvement schools yielded no significant interaction terms, a significant main effect of Initial Reading Band ($F = 13\cdot52$; $df = 2,98$; $P < 0\cdot001$), and a significant main effect of Project Group ($F = 6\cdot36$; $df = 1,98$; $P = 0\cdot013$).

A comparable analysis carried out on the data from the two extra teacher help schools yielded no significant interaction terms, and only one significant main effect, that of Initial Reading Band ($F = 15\cdot95$; $df = 2,78$; $P < 0\cdot001$).

The two sets of analyses reported so far suggest that the improvement in reading performance previously noted in children from the project's parent involvement groups has not "washed out" over time. Children from both schools and all initial reading bands who had taken part in the parent involvement exercise were reading better at 11 years of age than comparable children who had not. No such long term benefits were apparent in children who had received extra teacher help as part of the project.

### *(c) Reading performance in the three-year follow-up expressed in terms of the proportion of children falling into four score bands*

Following the same procedure as that adopted in the earlier study, the data for Schools 1 and 2 were pooled to give a combined parental help group and a combined control group. A similar pooling of groups was carried out for Schools 3 and 4. Table 2 gives the percentages of intervention and control children falling into four reading bands for the pooled data.

Chi-squared tests revealed that the distribution of children across categories was significantly different for the experimental and control children from the parent involvement schools ($\chi^2 = 8\cdot43$; $df = 3$; $P < 0\cdot05$), but not for the groups from the extra teacher help schools ($\chi^2 = 2\cdot73$; $df = 3$; NS). In the earlier evaluation report, it was noted that "parental help both reduced the proportion of failing readers (scores of 84 or less) and increased the proportion of able readers (scores of 115 or

TABLE 2

READING PERFORMANCE IN THE THREE-YEAR FOLLOW-UP, EXPRESSED IN TERMS OF THE PROPORTION OF CHILDREN FALLING INTO FOUR SCORE BANDS

| | Standardised score on the London Reading Test of: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 84 or less | | 85-99 | | 100-114 | | 115+ | | N |
| | % | (N) | % | (N) | % | (N) | % | (N) | |
| Schools 1 and 2 | | | | | | | | | |
| Parent involvement | 9·8% | (4) | 26·8% | (11) | 51·2% | (21) | 12·2% | (5) | 41 |
| Controls | 26·1% | (18) | 37·7% | (26) | 29·0% | (20) | 7·2% | (5) | 69 |
| | | | $x^2 = 8\cdot43$ with 3 df. $P < 0\cdot05$ | | | | | | |
| Schools 3 and 4 | | | | | | | | | |
| Extra teacher help | 17·6% | (6) | 29·4% | (10) | 41·2% | (14) | 11·8% | (4) | 34 |
| Controls | 12·5% | (7) | 46·4% | (26) | 33·9% | (19) | 7·1% | (4) | 56 |
| | | | $x^2 = 2\cdot73$ with 3 df. NS | | | | | | |
| Standardisation sample (Source: ILEA, 1980) | 16·7% | | 34·6% | | 33·6% | | 15·1% | | 5000 |

more). The lack of significant effect for the extra teacher help children appears most evident in the lower attainment band'' (p. 9). Exactly the same patterns are found in the follow-up data.

*(d) The effect of sample loss on the follow-up reading scores*

In order to assess whether or not distortions had been introduced into the figures as a result of differential sample loss, the reading scores obtained at the end of the project by the 48 children missing from the follow-up were compared with equivalent scores from children for whom follow-up data were available.

TABLE 3

NFER 'A' STANDARDISED SCORES OF FIRST YEAR JUNIORS WHO WERE, OR WERE NOT, SUBSEQUENTLY TRACED AND INCLUDED IN THE THREE-YEAR FOLLOW-UP

| School | Group | Included in follow-up Mean | SD | N | Missing from follow-up Mean | SD | N |
|---|---|---|---|---|---|---|---|
| 1 | Parent involvement | 106·6 | 9·18 | 18 | 108·6 | 4·67 | 5 |
| | Control | 95·8 | 11·86 | 40 | 94·8 | 5·56 | 9 |
| 2 | Parent involvement | 101·3 | 13·41 | 23 | 104·8 | 9·65 | 5 |
| | Control | 95·8 | 11·06 | 29 | 87·6 | 8·91 | 8 |
| 3 | Extra teacher help | 97·1 | 18·53 | 14 | 84·0 | 11·31 | 6 |
| | Control | 97·5 | 14·05 | 38 | 99·7 | 13·19 | 7 |
| 4 | Extra teacher help | 102·7 | 16·27 | 20 | 110·8 | 13·68 | 5 |
| | Control | 98·8 | 12·72 | 18 | 96·3 | 19·86 | 3 |

Table 3 summarises these findings. In School 1, loss from the Parent Involvement Group is likely to have had the effect of slightly lowering follow-up reading performance, whereas loss from the controls will probably have had a small effect in the opposite direction. The picture is the same in School 2, with above-average readers being lost from the Parent Involvement Group, and considerably below-average readers being lost from the Controls. In School 3, the opposite effect is apparent: a group of very poor readers were lost from the Extra Teacher Help Group, while loss from the Controls was mainly of relatively good readers. Finally, School 4 follows the same pattern as Schools 1 and 2, with differential sample loss acting to lower the performance of Extra Teacher Help Children, and raise that of their controls.

These findings suggest that the continued benefit apparent in children from the parent involvement groups is certainly not a statistical artefact of differential leaving patterns; if anything, the pattern of sample loss observed will have led to an underestimate of the benefit derived. The pattern is less clear in the extra teacher help schools, and it may well be the case that lasting benefits did accrue to children in School 4; no sign of benefit is apparent however in School 3.

## DISCUSSION

The data suggest that the Haringey Reading Project had a significant impact on children's reading careers. Children who had taken part in the parent involvement exercise were, as 11-year-olds, not just reading better than local controls but in some respects their performance also compared favourably with national standards: 63·4 per cent of the parent involvement children were reading at or above age level standard, compared to 36·2 per cent of within-school controls, and 48·7 per cent of children in the test standardisation sample.

At one extreme of the reading score distribution, the proportion of very weak readers found in the parent involvement group was reduced compared to both local and national levels; at the other extreme, the proportion of very good readers compared favourably with local controls, but did not reach the level found in the national sample. These differential results are consistent with the origins of the Haringey Project, which was designed to extend to all children a benefit — parental help — that a sizeable proportion of them would probably have received anyway. In these circumstances, the intervention would be predicted to draw in the tail of the schools' reading distribution, without affecting to the same degree performance at the upper end of the scale. Both parts of this prediction are supported by the results described.

No statistically significant long-term effects of the study's "extra teacher help" intervention were found in the follow-up data. The interpretation of findings from this part of the project was discussed in some detail in the first report; for present purposes, it is sufficient to note that the negative results are consistent with other long-term evaluations of school based special reading interventions (Hewison, 1982).

Comment is required on the slightly different patterns of results observed in the one-year and three-year follow-up exercises, the enduring effectiveness of the parent involvement project being more generally apparent in the latter than the former. Much the most likely explanation here relates to the reading tests used on the two occasions. To obtain a high score on the NFER "BD" test, used for the one-year follow-up, quite a lot of general knowledge (for example, about the French Revolution, or the destruction of Pompeii), as well as reading ability, seems to be required; whereas this is much less true of the London Reading Test, used subsequently. In middle class areas, children probably acquire general knowledge at a pace consistent with the development of their reading skills; in the project, although the acquisition of reading skills was speeded up, there is no reason why growth in general knowledge should have increased at the same rate. This potential "uncoupling" of the two areas of skill growth can make the interpretation of a test like the NFER "BD" quite problematical when it is used — as in the study — on children who may indeed be fairly good readers, but who come from backgrounds very different from those of most good readers in the test standardisation sample. It is not, of course, being argued here that growth in general knowledge is unimportant, only that the aim of the Haringey project was to improve reading performance, and its success should be judged accordingly.

The results presented in this and the first Haringey report show that the children who took part in the parent involvement exercise became — and remained — better readers. Why this happened is still unclear. At a theoretical level, the respective roles of motivation and practice are not known, nor is the relationship understood between the mechanisms at work in Haringey and those believed to be operating in behaviourally inspired approaches to parent involvement, such as Paired Reading (Morgan and Lyon, 1979; Topping, 1985).

At a more practical level, it is not known which "ingredients" of the intervention package mounted in Haringey were responsible for the project's success. School staff decided what reading material should be sent home, how its difficulty level should relate to the child's current level of performance, and how much checking of home reading should be carried out. A school which sent home, for example, much easier reading material than was chosen in the project might find that this altered the thrust of the intervention in important ways; certainly parents' and children's perceptions of the seriousness of the enterprise would be different in these circumstances, and it might be that the learning opportunities afforded to the children were also very different. In Haringey, researchers visited families at home to monitor the project's reception, and to offer very simple advice and guidance to parents about helping their children with reading; they also visited the schools, and discussed practicalities and progress with heads and class teachers. It may be that the motivation of children, families and school staff to persevere with the home reading scheme was maintained at unusually high levels by these means. It may also be the case that the very personal and supportive approach adopted by the researchers — in an area of considerable social disadvantage — had an impact on some families' beliefs and behaviour that was wider and more enduring than anything originally envisaged. Many factors, acting directly, and in interaction with one another, are likely to have contributed to the success of the Haringey Project, and in the present limited state of knowledge it is premature to identify any one component as the key to the project's success.

Two recently reported studies (Hannon, 1987; Tizard, 1987), using somewhat different mixtures of ingredients from the Haringey Project, have failed to identify reading benefits resulting from parental involvement; both authors have speculated on why this should be so. Further research will undoubtedly be required before the reasons for these discrepant findings are understood.

ACKNOWLEDGMENTS.—The author would like to offer her thanks to Mr Tony Lenney, former Director of Education in the London Borough of Haringey, for his permission to use the data reported here, and to his staff for their help in locating and abstracting that data. She would also like to thank Robert Drewett for his helpful comments on the manuscript of this paper, though responsibility for any errors it may contain must of course remain with her.

Correspondence and requests for reprints should be addressed to Dr. Jenny Hewison, Department of Psychology, University of Leeds, Leeds, LS2 9JT, England.

## REFERENCES

BENNETT. N. (1984). The London Reading Test. (Review). In LEVY, P., and GOLDSTEIN, H. (Eds.), *Tests in Education*. London: Academic Press.

GITTELMAN, R. (1985). Controlled trials of remedial approaches to reading disability. *J. Child Psychol. Psychiat.*, 26, 843-846.

HANNON, P. (1987). A study of the effects of parental involvement in the teaching of reading on children's reading test performance. *Br. J. educ. Psychol.*, 57, 56-72.

HEWISON, J. (1982). The current status of remedial interventions in children with reading problems. *Develop. Med. Child Neurol.*, 24, 183-185.

HEWISON, J. and TIZARD, J. (1980). Parental involvement and reading attainment. *Br. J. educ. Psychol.*, 50, 209-215.

ILEA (1980). *The London Reading Test*. London: Inner London Education Authority.

MORGAN, R., and LYON, E. (1979). Paired reading — a preliminary report on a technique for parental tuition of reading-retarded children. *J. Child Psychol. Psychiat.*, 20, 151-160.

TIZARD, B. (1987). Is parental help with reading effective? *Times Educational Supplement*, 3rd April.

TIZARD, J., SCHOFIELD, W. N., and HEWISON, J. (1982). Collaboration between teachers and parents in assisting children's reading. *Br. J. educ. Psychol.*, 52, 1-15.

TOPPING. K. (1985). An introduction to paired reading. In TOPPING, K., and WOLFENDALE, S. (Eds.), *Parental Involvement in Children's Reading*. Beckenham: Croom Helm.

(*Manuscript received 5th October*, 1987)

*Br. J. educ. Psychol.*, 58, 191-200, 1988

# REMEMBERING WRITTEN INSTRUCTIONS: TAB A GOES INTO SLOT C, OR DOES IT?

By JOHN A. GLOVER
*(Teachers College, Ball State University, Indiana, USA)*
ANNE L. HARVEY
*(Educational Testing Service)*
AND ALICE J. CORKILL
*(University of Nebraska, USA)*

Summary. Three experiments were performed in an attempt to determine if students' memory for written directions could be facilitated. Experiment 1 contrasted a control condition with a condition in which students expected a free recall post-test and a condition in which students expected to have to recall the directions in correct order. The results indicated that both experimental procedures resulted in significantly greater levels of overall recall than the control procedures. However, only the condition in which students expected a test of the order of information facilitated the recall of directions in order. Experiment 2 contrasted the learning of directions in a mnemonic ("one is a bun") condition, a paraphrase condition, and a control condition in which subjects expected a test of their ability to recall information in correct order. The results indicated that in both the paraphrasing and the mnemonic conditions students had significantly more total recall than in the control condition. Further, the students in mnemonic condition recalled significantly more directions in correct order than those in the paraphrase or the control conditions. Students in the paraphrase condition, however, did recall significantly more directions in correct order than students in the control condition. Experiment 3 replicated Experiment 2 but had students actually perform the task from memory. The results confirmed those of Experiment 2. Overall, the results suggest that students' memory for directions can be enhanced through various encoding manipulations. The results seem directly applicable to educational settings.

## INTRODUCTION

Most of us who have spent Christmas Eve assembling toys have had the frustrating experience of trying to remember the order in which we read instructions as we struggled with how to fit tab A into slot C. Similarly, even though written directions are usually available, students are sometimes frustrated when they attempt to recall sets of instructions they read for the completion of a procedure. For example, biology students may be given a description for how to prepare slides, chemistry students may be given the steps for the assembly of a titration apparatus, business students may be furnished with directions for operating a printer, and auto-mechanics students may be provided with a list of steps to follow in rebuilding wheel cylinders.

In each instance, the tasks could be completed by reading an instruction, attempting to complete it, and re-reading the instruction if it is forgotten or partially forgotten. However, it also seems that methods for facilitating students' recall of directions would prove beneficial by reducing the effort involved in completing tasks, the time necessary to learn the steps involved in a task, and the frustration experienced by students as instructions are forgotten.

Although a massive body of research on memory for prose has been generated over the past several years, little, if any, work has been devoted to techniques designed to facilitate students' memory for written directions. Logically, much of what we know about facilitating memory for prose should be applicable to aiding students' memory for written directions, but directions differ from most other forms of prose, because the order in which they are recalled is often of critical importance in completing tasks. In rebuilding wheel cylinders, for instance, some steps must precede others or the task will be incorrectly performed. Further, a set of directions may have no inherent passage structure other than the chronological order in which the directions are to be completed.

To date, research on memory for the order of information has come largely from studies of temporal order phenomena in memory for word lists (e.g., Tulving and Madigan, 1970; Tzeng, 1976; Zacks *et al.*, 1984). Subjects encoding word lists clearly rely on temporal order information at the time of recall (Underwood, 1949; Winograd, 1968; Tulving and Madigan, 1970; Tzeng and Cotton, 1980; Bellezza, 1982). While it has often been argued that the encoding of temporal order information is automatic (e.g., Zimmerman and Underwood, 1968; Tulving and Madigan, 1970; Toglia and Kimble, 1976; Hasher and Zacks, 1979; McCormack, 1982) and therefore not open to manipulation, recent work has indicated that both practice and ability influence memory for temporal information (Zacks *et al.*, 1984).

In the Zacks *et al.* (1984) study, for example, students were assigned to word list learning conditions in which they either practised encoding for temporal order or simply practised encoding for free recall of the lists. As one would expect, free recall improved among those students given practice in encoding for free recall. In addition, however, recall of temporal order information also improved among those given practice in encoding for a temporal order post-test. Further, high ability students recalled more words and more temporal order information about word lists than lower ability students.

Certainly, students' memory for meaningful sets of instructions is different from their memory for arbitrarily chosen words in a list. However, given that practice for the recall of temporal order information results in enhanced abilities to recall such information (Zacks *et al.*, 1984), it seems reasonable that altering the expectations students have when studying directions (e.g., Gagné, 1977), having students employ a mnemonic designed to facilitate recall of information in order (e.g., the "one is bun" technique; Miller *et al.*, 1960), and combining altered expectations with paraphrasing, a technique that has been shown to enhance recall (e.g., Glover *et al.*, 1981), would all have the possibility of facilitating recall of directions in correct order.

The current study presents three experiments on students' memory for written directions. The first two examine students' ability to recall directions in writing, the last focuses on students' ability to recall directions while actually performing the tasks. Experiment 1 was based on Zacks *et al.*'s (1984) study with word lists and was designed to determine if varying subjects' post-test expectations would influence their recall of a brief set of directions. Experiment 2 then contrasted the effects of a mnemonic condition with a paraphrase condition and a control condition on students' memory for a brief set of directions. Experiment 3 replicated Experiment 2 but actually had high school students perform the tasks from memory.

## EXPERIMENT 1

### *Method*

*Sample and setting.* The students were 45 undergraduate volunteers participating for course credit. None had experience with printer operation. All testing took place individually in a conference room across a period of three weeks.

*Materials and apparatus.* Directions consisting of a set of 10 steps for preparing a daisy wheel printer were identified via a task description process (see Appendix A).

The steps were validated by having six students not otherwise involved in the study try them out. The apparatus was a TRS-80 210 daisy wheel printer.

*Procedure.* Volunteers were randomly assigned to conditions: control, free-recall, and temporal order. The experimenter identified the printer for each student prior to the experimental procedures and verbally labelled each relevant part. Students were told they would need to know the names of the parts and were encouraged to ask questions until they were satisfied. The experimenter then quizzed each student by pointing to various parts of the equipment (in a random order, not the order in which the parts were used in readying the printer for operation) until they could correctly identify each relevant part of the printer.

Students in the control condition were told that they would read a set of steps for operating the printer and that they were to rate each step (on a five-point scale provided them) for clarity. Those in the free recall condition were told that they would read a set of steps for operating the printer and that they would then be tested over their ability to recall as many of the steps as possible. Those in the temporal order condition were given similar instructions but were told their test results would depend on their ability to remember the steps in correct order. Each student then read and studied the steps at his or her own pace. A free recall post-test was administered after they had finished studying and they were instructed to recall as many steps as possible in correct order. They were also told, however, to remember any steps or parts of steps they could if they could not remember their order. The printer was present on the table where the post-test was administered.

### *Results and Discussion*

Students' protocols were scored by two raters on two dependent measures: number of steps recalled ($r = 0{\cdot}98$) and number of steps recalled in their correct position ($r = 0{\cdot}98$). Synonymous terms were scored as correct (e.g., "roller" for "platen") on both measures. Steps were scored as being in correct order only if they appeared in the same position (i.e., one through ten) as in the original set of directions.

The recall, and recall in correct order, data were analysed via a one-way multivariate analysis of variance (MANOVA), with number of steps ordered first. The global test was significant, $F = 15{\cdot}12$, $P < 0{\cdot}01$. As both the number of steps recalled and recall in correct order were of equal interest, univariate analyses were then performed on each of the independent variables. The results indicated a significant effect among conditions on total recall, $F(2,42) = 12{\cdot}05$, $P < 0{\cdot}01$, $MSe = 3{\cdot}95$, and on recall in correct order, $F(2,42) = 8{\cdot}80$, $P < 0{\cdot}01$, $MSe = 6{\cdot}49$.

The univariate analyses were then followed up via the Tukey procedure ($\alpha = 0{\cdot}05$). On the total recall variable, students in both the free-recall and the temporal order conditions recalled significantly more steps than did those in the control condition. No other contrast was significant (see Table 1). On the recall in correct

TABLE 1

RECALL SCORES IN EXPERIMENT 1

| Conditions | Total Recall M | Total Recall SD | Recall in Order M | Recall in Order SD |
|---|---|---|---|---|
| Control (N = 15) | 4·67 | 2·38 | 3·53 | 2·61 |
| Free-recall (N = 15) | 7·93 | 1·62 | 3·87 | 2·85 |
| Temporal order (N = 15) | 7·53 | 1·88 | 7·06 | 2·12 |

order variable, students in the temporal order condition recalled significantly more steps in correct order than did subjects in either the free-recall or the control conditions (see Table 1). No other contrast reached significance.

The results generally confirm those of Zacks *et al.* (1984), although they had students practise studying for temporal order while we merely altered their expectations for post-testing. The results seem to indicate the possibility of using other encoding manipulations to enhance recall of directions and recall in correct order. In order to test two encoding procedures, a second experiment was conducted.

## EXPERIMENT 2

### *Method*

*Sample, setting, materials, and apparatus.* Students were 48 undergraduate volunteers participating for course credit. The setting, materials, and apparatus were identical to Experiment 1.

*Procedure.* Students were randomly assigned to conditions: control, mnemonic, and paraphrase. As in Experiment 1, they were familiarised and tested over knowledge of the printer. The control group was given the set of directions for preparing the printer for operation and asked to read and study them carefully as they would be tested for their ability to recall as many steps as possible in correct order upon completion.

Each student in the mnemonic condition was given a handout containing the "one is a bun" mnemonic rhyme and asked to learn it for later use in remembering information (see Appendix B). They studied the rhyme until they could recite it flawlessly on two occasions, and then were given practice using the mnemonic to remember a list of 10 words. Finally, they were given the steps for preparing the printer for operation and were also told that they would be tested for their ability to remember the steps in correct order.

Students in the paraphrase condition had paraphrasing demonstrated for them and were told that putting things to learn into their own words often helped recall (see Appendix C). Then they practised paraphrasing by writing out their paraphrases in learning a brief four-line poem. Finally, they were given the set of directions for preparing the printer for operation and asked to learn them by paraphrasing each step as they studied. They were also told that a post-test over their ability to recall the information in correct order would follow their studying.

All students were given as much time as they wanted to study the set of steps. The experimenter, however, recorded each individual's study time as an index of efficiency. The post-test was given in the presence of the printer as soon as the students completed studying.

### *Results and Discussion*

As in Experiment 1, the recall and recall in correct order data were analysed in a one-way MANOVA. The global test was significant, $F = 19{\cdot}68$, $P < 0{\cdot}01$. Univariate analyses indicated a significant effect among conditions on total recall, $F\ (2,45) = 16{\cdot}95$, $P < 0{\cdot}01$, $MSe = 2{\cdot}78$, and on recall in correct order, $F\ (2,45) = 42{\cdot}42$, $P < 0{\cdot}01$, $MSe = 4{\cdot}23$.

The univariate analyses were followed up via the Tukey procedure ($\alpha = 0{\cdot}05$). On the total recall variable, students in the mnemonic condition and paraphrase condition each recalled significantly more steps than those in the control condition. No other contrast was significant (see Table 2). On the recall in correct order variable, those in the mnemonic condition recalled significantly more steps in

TABLE 2

RECALL AND STUDY TIME SCORES IN EXPERIMENT 2

| Conditions | Total Recall M | Total Recall SD | Recall in Order M | Recall in Order SD | Study Time M | Study Time SD |
|---|---|---|---|---|---|---|
| Control (N = 16) | 6·75 | 2·79 | 2·81 | 2·59 | 5·60 | 1·26 |
| Mnemonic (N = 16) | 9·81 | 0·40 | 9·50 | 0·82 | 3·42 | 0·61 |
| Paraphrase (N = 16) | 9·62 | 0·62 | 6·50 | 2·31 | 5·63 | 1·14 |

Note: Study time is in minutes.

correct order than those in either the paraphrase or the control conditions. In addition, students in the paraphrase condition recalled significantly more steps in correct order than those in the control condition (see Table 2).

The study time data were analysed in a separate one-way univariate analysis. The results revealed a significant difference among conditions F (2,45) = 23·63, $P < 0{\cdot}05$, MSe = 1·09. When this difference was followed up via the Tukey procedure ($\propto = 0{\cdot}05$), it was observed that the students in the mnemonic condition used significantly less study time than those in either the paraphrase or the control conditions. No other contrast was significant (see Table 2).

As effective as the "one is a bun" mnemonic seemed to be in Experiment 2, it would appear to be limited to short lists of directions. Unfortunately, not all sets of directions are as brief as those employed in Experiments 1 and 2. In order to determine if it was possible to facilitate the learning of a long list of directions, an additional experiment was conducted. In this experiment, 30 undergraduates attempted to learn a set of 21 steps for assembling a distillation apparatus. Fifteen of them employed paraphrasing as described in Experiment 2. The remainder were in control condition. The results verified those of Experiment 2. That is, students who paraphrased the directions recalled significantly more of the steps than control students.

To this point, the results of Experiments 1 and 2 seem to indicate that students' memory for directions can be influenced by manipulating the ways in which the directions are encoded. There is, however, a severe limitation on Experiments 1 and 2. That is, the dependent measures were based solely on paper and pencil free recall post-tests. In most educational settings, students are seldom asked to write down directions from memory. Instead, the directions are used in actually performing some task. In order to determine if the procedures employed in Experiment 2 would facilitate students' actual task performance, an additional experiment was performed. In addition, in order to enhance the ecological validity of these experiments, they were conducted with high school students in a school meeting.

## EXPERIMENT 3

### *Method*

*Sample and setting*. Participants were 51 tenth grade students (mean age 15 years 9 months) who had indicated they would be taking a typing course the following year but who had not yet learned to type. In addition, none had used a printer. All data were collected individually over a period of two weeks in an office/storage area between two classrooms in a midwestern, urban high school.

*Procedure*. The procedures were identical to Experiment 2 except for the post-test. In this instance, when the students finished studying, they were asked to actually follow the steps from memory in preparing the printer for operation.

### *Results and Discussion*

The students were observed by the experimenter and another rater during the post-test. Both observers watched the students perform their task, recorded the steps the students performed and the order in which they were performed. In addition, upon completion each student's task was scored as successfully or not successfully completed. Inter-rater reliability was 0·92 for steps completed, 0·88 for steps completed in order, and 1·00 for successful task completion.

The observational data were analysed in a one-way MANOVA and the global effect was significant, F = 18·59, P < 0·01. Univariate analyses revealed a significant difference among conditions on total steps performed, F (2,48) = 12·92, P < 0·01, MSe = 3·68, and on steps performed in correct order, F (2,48) = 38·43, P < 0·01, MSe = 3·11. The Tukey procedure (α = 0·05) indicated that on total steps performed students in the paraphrase and the mnemonic conditions performed significantly more steps than those in the control condition. No other contrast was significant (see Table 3). On steps performed in correct order, students in the mnemonic condition performed significantly more steps in correct order than those in either the paraphrase condition or the control condition. In addition, those in the paraphrase condition performed significantly more steps in correct order than those in the control condition (see Table 3).

The analysis of study time was again performed separately via a univariate analysis. The results, F (2,48) = 7·13, P < 0·01, MSe = 3·12, indicated a significant difference among conditions. The Tukey procedure (α = 0·05), indicated that students in the mnemonic condition used significantly less study time than those in either the control or the paraphrase conditions. No other contrast was significant (see Table 3).

In addition, the proportion of students in the three conditions successfully completing the task was contrasted via a chi square test. The results were significant, $\chi^2$ (2 df) = 17·40, P < 0·01. A follow-up indicated that the proportion of students in the mnemonic condition correctly completing the task (17/17) did not differ significantly from the proportion of students in the paraphrase condition successfully completing the task (16/17). The proportion of students in each of the two experimental conditions who successfully completed the task was significantly greater than the proportion of students successfully completing the task in the control condition (7/17), however.

TABLE 3

RESULTS OF EXPERIMENT 3

| Conditions | Total Steps Performed | | Steps in Order | | Study Time | |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD |
| Control (N = 17) | 6·94 | 2·99 | 4·24 | 2·19 | 7·26 | 1·65 |
| Mnemonic (N = 17) | 10·00 | 0·00 | 9·53 | 0·80 | 5·43 | 1·62 |
| Paraphrase (N = 17) | 9·65 | 1·46 | 6·59 | 1·97 | 7·54 | 1·99 |

Note: Study time is reported in minutes. Steps are reported in steps actually performed by students.

The results of Experiment 3 confirm those of Experiment 2 and extend them to high school students who actually had to perform tasks rather than recall sets of steps in writing. In addition, a brief experiment using the steps in assembling a distillation apparatus also supported the results of Experiment 3. In this instance, students who paraphrased the steps involved in assembling the apparatus performed significantly more of the steps during testing than did control condition students. Apparently, the way in which students encode information influences their ability to remember and carry out written directions.

## GENERAL DISCUSSION

The results of the current study can be summarised as follows. First, when students are told they will be tested for the order in which they recall directions, they recall more of the directions in correct order than when they are merely told to expect a free recall post-test or when they are asked to rate the clarity of the directions. The overall amount of recall, however, is similar when students expect a free recall post-test and when students expect a test over their ability to recall information in correct order.

Second, the "one is a bun" mnemonic is highly effective for enhancing the recall of short lists of directions both on written measures of recall and when students must actually perform the tasks from memory. Third, paraphrasing to-be-learned directions enhances their memorability on both written and performance measures of recall, for both short and longer lists.

Although lists of presumably meaningful directions are considerably different from lists of randomly ordered words, the results of Experiment 1 can be seen as a confirmation and extension of the work of Zacks *et al.* (1984). In their study, Zacks and her colleagues gave students practice in the encoding of temporal order information in word lists and found, in part, that this practice resulted in significantly greater levels of memory for temporal order information. In the current Experiment 1, students were not given practice but had their levels of overall recall and recall in correct order enhanced by being given appropriate encoding instructions. If recalling a set of directions in correct order can be seen as a form of the recall of temporal order information, then Zacks *et al.'s* (1984) conclusion that recall for temporal order is not necessarily automatic is confirmed.

The results of Experiments 2 and 3 indicated that the "one is a bun" mnemonic condition resulted in significantly greater recall in correct order than either the paraphrase or the control conditions. In terms of overall recall, the mnemonic condition yielded significantly higher scores than the control conditions in Experiments 2 and 3 and the paraphrase condition in Experiment 3. In Experiment 2, the levels of overall recall in the mnemonic and the paraphrase conditions were highly similar.

The results of Experiments 2 and 3 were not surprising, although we are unaware of any previous research that has attempted to employ mnemonic techniques in the encoding of directions. The "one is a bun" mnemonic was originally described as a list learning procedure (Miller *et al.*, 1960), and a set of directions is a list of steps to be performed in a certain order. More generally, the results confirm the beneficial effects of mnemonics on students' memory for textual information (e.g., McCormick and Levin, 1984).

In terms of the paraphrasing manipulation employed here, the results generally support those previously obtained in studies in which subjects paraphrased specific segments of prose (e.g., Glover *et al.*, 1981; Dinnel and Glover, 1985). That is, paraphrasing resulted in significantly greater overall recall and recall in correct order than control conditions in Experiments 2 and 3, suggesting that paraphrasing is a relatively effective means of enhancing students' memory for sets of directions. Paraphrasing does not, however, seem to be superior to the "one is a bun" mnemonic employed in Experiments 2 and 3.

When study time is considered, the "one is a bun" mnemonic seemed more efficient than either control or paraphrase procedures. That is, after the mnemonic had been learned, significantly less time was required by students in the mnemonic conditions to learn their lists of directions than students in either the control or paraphrase conditions. This savings in study time was accompanied by significantly greater recall in correct order and overall recall equal to or greater than what was observed in the paraphrase conditions of Experiments 2 and 3. Study times were highly similar in the paraphrase and control conditions in Experiments 2 and 3.

An observation about the high rate of correct task performance in Experiment 3 seems to bear mention. All of the students in the mnemonic condition in Experiment 3 completed their task correctly. And, while our observations were not designed to provide elapsed times per direction or to try to rate hesitations, it seemed that these students performed much as would have been expected from the free-recall results of Experiment 2. That is, they generally proceeded from direction one through direction 10 in a smooth, unhesitating fashion, much as would be expected if they were simply recalling a well-learned list of materials. In contrast, students in the paraphrase condition in Experiment 3 seemed to behave somewhat differently. Although 16/17 in the paraphrase condition of Experiment 3 actually did complete their tasks successfully, there seemed to be considerably more hesitation and trial and error in their performances, usually occurring over the last one-third or so of the directions. They were generally able to complete their tasks but it seemed that the high rate of task completion may have been due to their remembering enough to "get close" to task completion, allowing for a form of problem solving to complete the tasks. That such successful "problem solving" was seldom observed in the control conditions suggests that paraphrasing did indeed enhance subjects' memory for the directions but that other processes may have contributed to the likelihood of actual task completion. Our hypothesis about the differences between how students in the mnemonics and the paraphrase conditions completed their tasks in Experiment 3 are conjectural at this point, however, and will require further research.

The results of the current study seem to have direct implications for educational practice. Students' ability to remember directions can be enhanced through either the "one is a bun" mnemonic or paraphrasing procedures. Neither technique is difficult to implement and the benefits seem clear. In general, it would seem that the mnemonic is superior for brief sets of directions, while the paraphrase procedure is workable for both short and long lists of directions. It would appear that teachers ought to consider mnemonic or paraphrasing procedures to help their students remember lists of written directions.

The current study is limited in that we examined only two procedures for manipulating the encoding of written directions. Certainly, a review of the memory for prose literature indicates that there are many other potential methods for enhancing students' memory for written directions. Further, a review of the mnemonics literature also indicates several other possibilities for facilitating students' recall of directions. Finally, the current study was restricted to written directions and there is no assurance that the techniques described here may generalise to students' memory for oral directions.

Correspondence and requests for reprints should be addressed to Professor John A. Glover, Teachers College, Ball State University, Muncie, Indiana, USA 47306.

## REFERENCES

BELLEZZA, F. S. (1982). Updating memory using mnemonic devices. *Cogn. Psychol.*, 14, 301-327.

DINNEL, D., and GLOVER, J. A. (1985). Advance organizers: encoding manipulations. *J. educ. Psychol.*, 77, 5, 514-521.

GAGNE, R. M. (1977). *The Conditions of Learning* (3rd. ed.). New York: Holt, Rinehart, and Winston.

GLOVER, J. A., PLAKE, B. S., ROBERTS, B., ZIMMER, J. W., and PALMERE, M. (1981). Distinctiveness of encoding: the effects of paraphrasing and drawing inferences on memory for prose. *J. educ. Psychol.*, 73, 736-744.

HASHER, J., and ZACKS, R. T. (1979). Automatic and effortful processes in memory. *J. exp. Psychol.: gen.*, 108, 356-388.

MCCORMACK, P. D. (1982). Temporal coding and study-phase retrieval in young and elderly adults. *Bull. psychon. Soc.*, 20, 242-244.

MCCORMICK, C. B., and LEVIN, J. R. (1984). A comparison of different prose-learning variations of the mnemonic keyword method. *Am. educ. Res. J.*, 21, 379-398.

MILLER, G. A., GALANTER, E., and PRIBRAM, K. H. (1960). *Plans and the Structure of Behavior.* New York: Holt, Rinehart, and Winston.

TOGLIA, M. P., and KIMBLE, G. A. (1976). Recall and use of serial position information. *J. exp. Psychol.: hum. Learn. Mem.*, 2, 431-445.

TULVING, E., and MADIGAN, S. A. (1970). Memory and verbal learning. *Ann. Rev. Psychol.*, 21, 437-484.

TZENG, O. (1976). A precedence effect in the processing of verbal information. *Am. J. Psychol.*, 89, 577-599.

TZENG, O., and COTTON, B. (1980). A study-phase retrieval model of temporal coding. *J. exp. Psychol.: hum. Learn. Mem.*, 6, 707-716.

UNDERWOOD, B. J. (1949). Proactive inhibition as a function of time and degree of prior learning. *J. exp. Psychol*, 39, 24-34.

WINOGRAD, E. (1968). List differentiation as a function of frequency and retention interval. *J. exp. Psychol. Monogr. Suppl.*, 76, (2 pt. 2), 1-18.

ZACKS, R. T., HASHER, L., ALBA, J. W., SANFT, H., and ROSE, K. C. (1984). Is temporal order encoded automatically? *Mem Cognit.*, 12, 387-394.

ZIMMERMAN, J., and UNDERWOOD, B. J. (1968). Ordinal position knowledge within and across lists as a function of instructions in free-recall learning. *J. gen. Psychol.*, 79, 301-307.

(*Manuscript received 17th September,* 1986)

## APPENDIX A

### STEPS FOR OPERATING THE PRINTER

1. Open clamps on side of printer.
2. Put end of paper over the sprockets.
3. Insert sprockets into holes in paper.
4. Close clamps.
5. Turn platen knob until paper appears in front.
6. Set dip switch to "feed".
7. Set line switch to "on line".
8. Set pitch switch to "ps".
9. Turn on power switch.
10. Push self test switch.

## APPENDIX B

### DIRECTIONS AND RHYME FURNISHED SUBJECTS IN MNEMONIC CONDITIONS

Rhyming mnemoics can be helpful in learning lists of items or steps in a process. Below is a rhyme you probably are already familiar with. Even so, take the time to study it now. You will have to be able to recite this rhyme to the experimenter flawlessly on two occasions before you can proceed further in today's activity. You will use this rhyme to help you remember some information.

One is a bun
Two is a shoe
Three is a tree
Four is a door
Five is a hive
Six is sticks
Seven is heaven
Eight is a gate
Nine is a pine
Ten is a hen

Now that you have learned the rhyme, let's see how it can be used to help you remember lists of items. To use the rhyme imagine the first item in a list interacting in your mind with the image of a bun. For example, let's suppose that the first item in a list of things to be purchased at a hardware store is a hammer.

You could imagine a hammer smashing a bun flat for your first image. After the image is clear in your mind, you move to the next item in the list and imagine it interacting with a shoe. If the second thing you were to purchase at the hardware was a pound of nails, you could imagine a shoe filled with nails. This process is then repeated until all the list of things to remember is learned. Then, when you are to recall the items, you go back through the rhyme and at each line remember the image the words call up (e.g., one is a bun should evoke the image of a hammer smashing a bun). Practise your use of the technique on the grocery list below:

apples
bread
milk
butter
bologna
tomatoes
bacon
peaches
tuna
pickles

## APPENDIX C

### DIRECTIONS FURNISHED SUBJECTS IN PARAPHRASE CONDITIONS

One technique that often helps people commit information to memory is paraphrasing. For example, if you wished to recall the meaning of the sentence, "The fat man slept by the fire," you could take a moment and paraphrase the sentence into, "The corpulent male snoozed near the conflagration." The act of putting information into your own words seems to strengthen the memory trace we have for the meaning of materials.

Below is a brief four line poem. Please paraphrase it in the space provided. Later, you will be asked to paraphrase some materials in order to help you remember them.

The curfew tolls the knell of parting day,
The lowing herd winds slowly o'er the lea,
The plowman homeward plods his weary way,
And leaves the world to darkness and to me.

*Br. J. educ. Psychol.*, 58, 201-204, 1988

# SCIENCE AND HANDEDNESS

BY T. J. HARVEY
*(School of Education, University of Bath)*

SUMMARY. The aims of this study were (a) to investigate whether left-handedness is related to intelligence, and (b) as a consequence of investigating the incidence of left-handedness whether science can be classified as a one-side of the brain activity. To test for the relationship between left-handedness and intelligence, 800 fifth formers were given the Oldfield Handedness Inventory, as were a 1,000 sixth formers. The test was also given to 400 university science undergraduates. The results indicated (a) that left-handedness did not appear to be a function of intelligence, (b) for males science could possibly be considered as a right-hand side of the brain activity, and (c) for females science is possibly an across-the-brain activity.

## INTRODUCTION

THE vast majority of the human population have a dominant hand and few people can be considered to be truly ambidextrous. This dominance is a consequence of the asymmetry of the brain. Bakan (1971) claims that everyone may be classified as a right or left hemisphere brain person, depending upon which hemisphere guides the bulk of an individual's behaviour.

A substantial amount of work on split-brain patients appears to indicate that the left hemisphere is important for analytic processes, especially the production and understanding of language, whereas the right hemisphere is responsible for spatial skills and musical ability. Hemispheric specialisation is not an all-or-none phenomenon but rather falls on a continuum. Each half of the brain appears capable of perceiving, learning and remembering independently of the other, but some differences exist in the way, and in the efficiency, with which each half deals with the incoming information.

Orstein (1978) suggests that lawyers and artists use different halves of the brain in their work and that the differences between the halves show up in activities not related to their work. If spatial ability is mainly related to the right hemisphere, and since it is this hemisphere which controls left-handedness, one might expect a greater incidence of left-handedness in groups such as architects who rely upon spatial ability. Such an expectation was confirmed by Peterson and Lansky (1974).

Lansdell (1962) and McGlone (1978) showed that females were more able to operate between the hemispheres than males, i.e., their brains were less lateralised than for males. This work was further supported by Wade *et al.* (1975) when comparing the ratios of the length of the right temporal brain to the left temporal brain for males and females at post-mortems. Overall he found that this ratio was invariably less than 1 for males, but frequently exceeded 1 for females. A much further discussion of gender and asymmetry is given by Springer and Deutsch (1985).

The aim of this present study is to investigate whether left-handedness is in any way related to intelligence, and to see whether science can be classified as a one side of the brain activity.

## METHOD

The test for handedness was based upon the Oldfield Handedness Inventory (1971) which generates a laterality quotient ranging from +100 (total right-handedness in all tasks) to $-100$ (complete left-handedness). People with positive laterality quotients were thus categorised as right-handed and those with negative quotients as left-handed.

The Oldfield Inventory was administered to 838 fifth year pupils in several comprehensive schools in the South West of England. All the fifth year pupils in the schools were tested and so it is reasonable to assume that they represent, certainly within this area, the full cross-section of ability and intelligence. The results of the survey are given in Table 1.

TABLE 1

NUMBERS OF RIGHT- AND LEFT-HANDED FIFTH YEAR PUPILS AND THE RATIO LEFT- TO RIGHT-HANDED

| | Male | | Female | |
|---|---|---|---|---|
| | Numbers | Ratio | Numbers | Ratio |
| Left-handed | 53 | 1 | 52 | 1 |
| Right-handed | 354 | 6·7 | 379 | 7·3 |

It has been suggested by Oldfield (1971) that in some studies but not in others there is a higher incidence of left-handedness amongst males than females. A $\chi^2$ test applied to the results of Table 1 shows that no differences ($P > 0 \cdot 1$) exist between males and females. Further analysis shows that the ratios for the sexes is the same even if the results are further subdivided such that only strongly left-handed groups (laterality quotients) between $-50$ and $-100$ are considered.

The inventory was then administered to 1,000 sixth form pupils in local comprehensive schools, the assumption being that since they were sixth formers, they were as a group, on average, of higher intelligence than the groups of fifth formers.

## RESULTS

The results of this survey are given in Table 2.

If one compares the results of Table 1 and Table 2, to see if the incidence of left-handedness is greater in the sixth form population, a $\chi^2$ test shows that no significant difference occurs ($P > 0 \cdot 1$). This would imply that left-handedness is not a function of intelligence. There are obvious limitations upon this experimental design. No attempt has been made to quantify the intelligence levels of the fifth form and sixth form groups, other than to assume that by their very nature the sixth formers have a higher average intelligence. It is interesting to note that anecdotal evidence from many school teachers suggests that there is an unusually high

TABLE 2

NUMBERS OF RIGHT- AND LEFT-HANDED SIXTH YEAR PUPILS AND THE RATIO LEFT- TO RIGHT-HANDED

| | Male | | Female | |
|---|---|---|---|---|
| | Numbers | Ratio | Numbers | Ratio |
| Left-handed | 61 | 1 | 50 | 1 |
| Right-handed | 489 | 8·0 | 400 | 8·0 |

incidence of left-handedness in the remedial/lower ability groups. To test fully any hypothesis relating left-handedness to intelligence would require an extremely large sample in which possibly the verbal reasoning quotient of each person would have to be measured. However, for the purpose of this study it is sufficient to conclude that no significant difference between the incidence of left-handedness of the two groups was found, i.e., it does not appear to be a function of intelligence, or at least not strongly related to intelligence.

The Inventory was then administered to approximately 400 university students at present studying undergraduate courses in science. The results of this survey are given in Table 3.

TABLE 3

NUMBERS OF RIGHT- AND LEFT-HANDED UNIVERSITY SCIENCE STUDENTS AND THE RATIO LEFT- TO RIGHT-HANDED

| | Male | | Female | |
|---|---|---|---|---|
| | Numbers | Ratio | Numbers | Ratio |
| Left-handed | 48 | 1 | 14 | 1 |
| Right-handed | 232 | 4·8 | 104 | 7·4 |

These results are to be compared with those of Table 2, for it is from a similar population that these students were presumably initially selected prior to going to university. Clearly there is no difference in the incidence of left-handedness amongst the females, but if a $\chi^2$ test is applied to the males there is a significant difference ($P < 0 \cdot 02$) between the two groups. The incidence of left-handedness amongst the males has almost doubled for the university science students compared with the male sixth formers.

## DISCUSSION

If we take the findings from the first part of the study that left-handedness is unlikely to depend upon intelligence, this difference cannot be explained by saying that it is because only the more intelligent of the sixth formers go on to university. In any case it would appear strange to apply an intelligence argument to explain the difference for the males but then not require to apply it for the females.

A more realistic explanation is to suggest that science for males is likely to be a right-hand side brain activity, i.e., giving consequently a higher incidence of left-handedness. For the females it appears to support the view that science is an across the brain activity, which supports the earlier view that the brain of the female is less lateralised than for the male.

This work is now being extended to university mathematics and language students, for it could give an interesting insight into the difference in laterality between males and females.

Correspondence and requests for reprints should be addressed to Dr. T. J. Harvey, School of Education, University of Bath, Claverton Down, Bath, BA2 7AY, England.

## REFERENCES

BAKAN, P. (1971). The eyes have it. *Psychol. Today*, 4, 64-69.

LANSDELL, H. (1962). A sex difference in effect of temporal lobe surgery on design preference. *Nature*, 194, 852-854.

McGlone, J. (1978). Sex difference in functional brain asymmetry. *Cortex*, 14, 122-128.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychol.*, 9, 97-113.

Orstein, R. (1978). The split and whole brain. *Hum. Nature*, 1, 78-83.

Peterson, J. M., and Lansky, L. M. (1974). Left-handedness among architects: some facts and speculation. *Percept. mot. Skills*, 38, 547-550.

Springer, S., and Deutsch, G. (1985). *Left Brain, Right Brain.* New York: W. H. Freeman.

Wade, J. A., Clarke, R., and Hamm, A. (1975). Cerebral hemisphere asymmetry in humans. *Archs. Neurol.*, 32, 239-246.

(*Manuscript received 20th July*, 1987)

# RESEARCH NOTES

*Br. J. educ. Psychol.*, 58, 205-211, 1988

# LARGE VERBAL — NON-VERBAL ABILITY DIFFERENCES AND UNDERACHIEVEMENT

By JOYCE WHITTINGTON

SUMMARY. In a large-scale study (over 10,000 11-year-olds) of verbal-non-verbal ability differences and underachievement it was found that children who had large V-NV ability differences were more likely to be underachieving in reading comprehension or mathematics or both than children of the same overall ability with more similar V and NV components. The rates for "severe" underachievement (2 standard errors below expectation, based on regression of achievement scores on ability scores) were two to ten times greater among the large V-NV ability differences group of children, nearly 20 per cent of whom were severely underachieving in reading comprehension, as were over 5 per cent in mathematics.

## INTRODUCTION

Large differences between verbal (V) and non-verbal (NV) abilities are known to occur together with underachievement in reading or mathematics in some cases of specific learning disability and in some congenital conditions. For example, children with specific language difficulties such as dyslexia often have depressed verbal ability scores (Kinsbourne and Warrington, 1963a; Crosby, 1976); some children who exhibit Gerstmann syndrome signs experience mathematical difficulties and depressed non-verbal ability scores (Kinsbourne and Warrington, 1963b; Benson and Geschwind, 1970); and females with a missing x-chromosome also show this latter pattern (Money, 1964; Reske-Nielsen *et al.*, 1982).

These observations lead one to ask: are all large verbal-non-verbal differences associated with learning difficulties and, conversely, are all learning difficulties associated with large verbal-non-verbal differences?

It is well-known that not all cases of learning difficulties or of underachievement are associated with large differences between verbal and non-verbal abilities (even in the studies quoted above). This has been variously interpreted. One extreme view is that large V-NV differences occur naturally and have no educational implications, so that large V-NV differences and underachievement co-occur by chance. The opposite extreme holds that all large V-NV differences are pathological and indicate abnormal brain function (actual damage, malformation, abnormal maturational development, etc.) which leads to underachievement, and that underachievement in the absence of large V-NV differences is caused by different factors. The first of these extreme positions would imply that there would be no differences in achievement between groups whose members have large V-NV ability differences and groups whose members have no large differences in abilities. The other extreme position would imply that all large V-NV ability differences may be expected to result in lower achievement levels in at least some areas.

Discussing the significance of large verbal-non-verbal differences, Berk (1982) distinguished between abnormality, statistical significance, and diagnostic significance of large V-NV differences and concluded that there is no proven diagnostic significance in large V-NV differences. He noted, however, that such discrepancy scores are often used as indicative of brain abnormality in some clinical diagnostic procedures. His concern was a possible connection between large V-NV differences and neurological status, and he did not report specifically on connections between V-NV differences and achievement variables.

A search of the literature has failed to reveal any large-scale investigation of the relationship between large V-NV differences and underachievement such as we shall describe. Our investigation concerns the relationship between very large verbal-non-verbal ability differences and underachievement in a national sample of 11-year-old schoolchildren.

## METHOD

### *The children*

The children were all those born in one week of March 1958, in England, Scotland and Wales, for whom relevant data* were available. At birth they were the subjects of the British Perinatal Mortality Survey (Butler and Bonham, 1963). They were followed up in 1965 at the age of 7 (Davie *et al.*, 1972), at the age of 11 (Davie, 1973), and at 16 years (Fogelman, 1976) as the National Child Development Study (NCDS). Medical questionnaires were completed for all these stages (i.e., at birth and ages 7, 11 and 16) and parents completed questionnaires containing social and behavioural data; schools and teachers completed educational questionnaires containing educational and behavioural assessments at 7, 11 and 16 years, and also ability tests, constructed by the NFER and standardised for use with 11-year-olds, at 11 years; and the children completed individual questionnaires at age 16.

We have used the results of the verbal and non-verbal ability tests and the mathematics and reading achievement tests at age 11 to investigate the relationship between large V-NV differences in ability and underachievement in mathematics and reading.

### *The tests*

The verbal and non-verbal abilities tests each consisted of 40 multiple-choice questions (examples of which are given in Fogelman, 1976).

These were combined into a single test instrument in which verbal and non-verbal items alternated. Three results were recorded: number of correct answers to verbal items, number of correct answers to non-verbal items, and total number of correct answers. The combined test had been constructed by the NFER for use in previous large-scale research (Douglas, 1964). It had a reliability (KR formula 20, 363 subjects) of 0·94, and correlated 0·93 with NFER Verbal Test 8A. Separate reliabilities and validities for the two subscales have never been published, and the NFER now have no records of their compilation, but presume that at the time of construction they were measurably reliable and valid (personal communication).

These verbal and non-verbal ability tests are clearly not directly comparable with the verbal and non-verbal subscales of larger test batteries. They obviously also have a large common component; in fact the verbal and non-verbal subtests were very similar in three important respects:

(a) The type of reasoning required.

(b) The layout of the test material on the page.

(c) Both were paper-and-pencil timed tests.

It is therefore not surprising that in our study they were found to correlate 0·81. However, an examination of the test material does suggest that whereas identification of word meanings and relationships is necessary for the verbal items, identification of shapes, their orientation and relationships is necessary for the non-verbal items. Evidence that the tests did, nevertheless, tap some different basic abilities comes from three sources: the fact that some children did score very much more highly on one scale than on the other, that among such large differences significantly more favoured non-verbal items, and that among such large differences there was a significant interaction between sex and direction of ability difference (see the Results section).

The reading comprehension and mathematics tests were specially constructed for use in the National Child Development Study by the NFER. The former was designed to be parallel to the Watts-Vernon Test, and had a reliability (KR formula 21, N = 300) of 0·82. The mathematics test contained both mechanical and problem items and tested a wide range of mathematical skills; there were 40 questions and the reliability (KR formula 21, N = 300) was 0·94.

### *The Difference Group*

The distributions of raw scores on the verbal and non-verbal tests were not quite normal (especially that of the verbal test), had small floor and ceiling effects, and were not directly comparable. The scores have therefore been converted into percentiles, and discrepancies are discussed in terms of percentile differences.

---

* Relevant to the present analysis means that ability and achievement scores at 11 years were available. Details of missing data can be found in Davie *et al.* (1972), Butler and Bonham (1963) and Fogelman (1976).

Following this convention, all children with V-NV ability differences of 50 percentiles or more were assigned to our difference (D) group. The latter comprised 147 children of whom 99 were boys, and 34 were girls, while the sex of 14 children was not recorded.

### *Underachievement*

Following Thorndike (1963) and Rutter and Yule (1975), underachievement was defined in terms of the regression of achievement variables on total ability scores. The regression line was taken to predict a child's achievement, given his total ability score. Achievement scores below prediction were classed as underachievement, those above prediction as overachievement. Of course, the significance of any under- or over-achievement depends on an arbitrary cut-off, as described by Thorndike who advocated 2 standard errors.

We shall initially ignore the question of significance when we look at underachievement and begin by comparing the distributions of the actual achievement scores with those predicted by the total ability scores of our difference group. We shall then take a cut-off of 2 standard errors below prediction to define "severe underachievement".

## RESULTS

### *Prevalence of large V-NV ability differences*

The distribution of the V-NV ability percentile differences for the whole population is shown in Table 1. The distribution of these V-NV discrepancies is not normal, having too many subjects with very large discrepancies. It is, of course, possible that this result has arisen as a statistical artefact, but if not it suggests that large V-NV differences may be abnormal in a psychologically meaningful way as well as statistically.

TABLE 1

DISTRIBUTION OF V-NV ABILITY DIFFERENCES IN THE POPULATION OF 11-YEAR-OLDS

| Percentile score difference | 0 to 10 | 10 to 20 | 20 to 30 | 30 to 40 | 40 to 50 | 50 to 60 | 60 to 70 | 70 to 80 | 80 to 90 |
|---|---|---|---|---|---|---|---|---|---|
| Percentage population V > NV | 25·0 | 13·0 | 7·1 | 3·1 | 1·3 | 0·40 | 0·11 | 0·02 | 0·0 |
| Percentage population NV > V | 25·0 | 13·0 | 6·9 | 2·9 | 1·3 | 0·62 | 0·20 | 0·07 | 0·01 |
| Percentage population Total | 50·0 | 26·0 | 14·0 | 6·0 | 2·6 | 1·0 | 0·31 | 0·09 | 0·01 |

### *The Difference Group*

Table 2 shows the distribution of V-NV ability differences in the D-group by direction of ability difference and by sex of child. There was a tendency for girls to have higher verbal scores, and for boys to have higher non-verbal scores, significantly so in the D-group ($\chi^2$ = 11·1, df = 2, P < 0·01), but very weakly in the whole population. There were significantly more boys than girls in the D-group ($\chi^2$ = 29·8, df = 1, P < 0·001). This non-random composition of the D-group suggests that large V-NV differences have arisen systematically and are not solely due to errors of measurement.

TABLE 2

DISTRIBUTION OF V-NV DIFFERENCES BY DIRECTION AND BY SEX IN THE D-GROUP

| | BOYS | GIRLS | ? |
|---|---|---|---|
| V > NV | 6 | 9 | 1 |
| NV > V | 93 | 25 | 13 |

*Underachievement*

We begin by looking at the numbers of children in the D-group who scored above and below prediction respectively on the two achievement tests, predictions being based on total ability scores.

Table 3 shows the distributions of above- and below-prediction scores for reading comprehension and for mathematics, for boys and girls separately. The figures also show (in brackets) whether verbal or non-verbal abilities were dominant.

TABLE 3

ACHIEVEMENT SCORES RELATIVE TO PREDICTION BY SEX AND BY DIRECTION OF V-NV DIFFERENCE FOR THE D-GROUP

| | Below prediction in both RC and M | Below prediction in RC only | Below prediction in M only | At or above prediction in both |
|---|---|---|---|---|
| Boys | 59 (all NV > V) | 13 (11 NV > V) | 14 (12 NV > V) | 13 (11 NV) |
| Girls | 23 (22 NV > V) | 2 (all V > NV) | 6 (2 NV > V) | 3 (1 NV > V) |
| ? | 10 (all NV > V) | 0 | 1 (NV > V) | 3 (2 NV > V) |

RC = Reading Comprehension M = Mathematics

From Table 3 it can be seen that 107 children out of 147 achieved below prediction in reading comprehension, while 113 children out of 147 achieved below prediction in mathematics. These proportions are significantly higher than those for the population as a whole ($\chi^2 = 29\cdot6$, df = 1, $P < 0\cdot001$ for reading and $\chi^2 = 41\cdot4$, df = 1, $P < 0\cdot001$ for mathematics).

More important, from the point of view of this article, than the incidence of scores above or below predicted values is the incidence of severe underachievement. Thorndike (1963) argued that underachievement should be considered significant if scores are two standard errors below predicted values, and this was the cut-off used by Rutter and Yule (1975) in their demonstration that specific reading retarded children differed significantly from reading backward children. Using this criterion of two standard errors below prediction, three of the D-group (2 per cent) were significantly underachieving in both reading comprehension and mathematics while a further five (3·4 per cent) were significantly underachieving specifically in mathematics, and 26 others (17·7 per cent) specifically in reading comprehension. These rates of 2, 3·4 and 17·7 per cent for the D-group compare with population rates of 0·002, 0·02 and 0·015 per cent respectively.

The children in our difference group excluded those of very low or very high total ability by definition. Because we imposed a 50 centile points difference between V and NV abilities as criterion for inclusion, only those children with total ability scores roughly between the 25th

TABLE 4

DISTRIBUTION OF SEVERE UNDERACHIEVEMENT BY SEX AND BY DIRECTION OF V-NV DIFFERENCE FOR THE D-GROUP

| | RC and M | M only | RC only |
|---|---|---|---|
| Boys | 1 (NV > V) | 1 (NV > V) | 18 (all NV > V) |
| Girls | 1 (NV > V) | 4 (3 NV > V) | 6 (all NV > V) |
| ? | 1 (NV > V) | 0 | 2 (all NV > V) |

TABLE 5

ACHIEVEMENT SCORES FOR SELECTED ABILITIES BROKEN DOWN BY V-NV DIFFERENCES

| | | | | | V > NV | | | | Percentile Differences | | | NV > V | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 70 to 80 | 60 to 70 | 50 to 60 | 40 to 50 | 30 to 40 | 20 to 30 | 10 to 20 | 0 to 10 | 0 to 10 | 10 to 20 | 20 to 30 | 30 to 40 | 40 to 50 | 50 to 60 | 60 to 70 | 70 to 80 | 80 to 90 |
| | RC | | | | | | | 8·2 | 8·4 | 6·9 | 6·6 | 0·0 | | | | | | |
| TA = 15 | M | | | | | | | 1·8 | 3·8 | 2·4 | 2·3 | 0·0 | | | | | | |
| | (N) | | | | | | | (9) | (64) | (18) | (9) | (1) | | | | | | |
| | RC | | | | | 10·5 | 13·1 | 11·1 | 11·2 | 9·9 | 9·4 | 8·3 | 11·2 | 5·8 | | 9·0 | | |
| TA = 25 | M | | | | | 5·8 | 7·3 | 6·3 | 6·8 | 7·4 | 6·3 | 4·7 | 8·8 | 3·3 | | 6·0 | | |
| | (N) | | | | | (6) | (13) | (29) | (49) | (25) | (40) | (7) | (6) | (4) | | (1) | | |
| | RC | | | 12·0 | 13·0 | 13·3 | 14·1 | 12·2 | 13·3 | 13·1 | 12·6 | 11·9 | 10·5 | 10·5 | 6·5 | 5·0 | 3·0 | |
| TA = 30 | M | | | 8·0 | 11·0 | 6·2 | 11·7 | 9·2 | 9·6 | 9·9 | 9·0 | 10·7 | 7·1 | 9·4 | 7·5 | 4·0 | 3·0 | |
| | (N) | | | (1) | (1) | (6) | (19) | (27) | (56) | (28) | (46) | (20) | (14) | (11) | (2) | (1) | (2) | |
| | RC | | | | 17·5 | 15·5 | 15·3 | 14·8 | 14·6 | 13·5 | 13·8 | 12·6 | 11·5 | 9·2 | 11·0 | | 9·0 | |
| TA = 35 | M | | | | 16·0 | 11·0 | 14·7 | 12·9 | 12·6 | 10·9 | 12·3 | 10·5 | 10·3 | 8·5 | 8·7 | | 3·0 | |
| | (N) | | | | (2) | (22) | (17) | (56) | (32) | (32) | (60) | (21) | (12) | (10) | (3) | | (1) | |
| | RC | | 16·0 | 21·0 | 15·0 | 15·1 | 15·8 | 15·8 | 15·8 | 15·5 | 15·2 | 15·2 | 13·6 | 11·3 | 12·3 | 11·5 | | |
| TA = 40 | M | | 14·0 | 16·0 | 15·4 | 14·0 | 14·0 | 15·9 | 16·0 | 15·6 | 14·8 | 14·0 | 12·6 | 13·3 | 9·8 | 18·0 | | |
| | (N) | | (1) | (1) | (9) | (10) | (15) | (63) | (34) | (31) | (29) | (33) | (19) | (4) | (4) | (2) | | |
| | RC | | | 16·8 | 15·3 | 17·3 | 16·9 | 16·4 | 17·1 | 15·4 | 15·9 | 15·5 | 13·1 | 12·3 | 14·3 | | | |
| TA = 43 | M | | | 19·2 | 10·0 | 17·5 | 17·0 | 17·2 | 17·8 | 15·0 | 15·8 | 15·6 | 10·6 | 14·1 | 12·0 | | | |
| | (N) | | | (5) | (3) | (18) | (14) | (25) | (33) | (50) | (46) | (17) | (8) | (9) | (3) | | | |
| | RC | | | 23·5 | 20·3 | 18·6 | 18·5 | 17·1 | 17·0 | 16·8 | 15·0 | 15·1 | 14·2 | 14·1 | 13·5 | 18·0 | | |
| TA = 45 | M | | | 22·5 | 16·0 | 17·1 | 18·0 | 16·0 | 19·1 | 15·9 | 15·7 | 16·0 | 15·3 | 10·5 | 19·0 | 18·0 | | |
| | (N) | | | (2) | (7) | (11) | (44) | (34) | (52) | (44) | (39) | (41) | (18) | (8) | (2) | (1) | | |
| | RC | | 16·0 | | 17·4 | 18·9 | 19·4 | 18·4 | 18·0 | 19·2 | 17·6 | 17·2 | 14·9 | 16·0 | 16·8 | | | |
| TA = 50 | M | | 20·5 | | 21·1 | 20·3 | 20·4 | 21·8 | 19·3 | 22·5 | 20·2 | 21·8 | 15·6 | 20·9 | 16·6 | | | |
| | (N) | | (2) | | (7) | (27) | (20) | (37) | (32) | (43) | (76) | (18) | (8) | (8) | (5) | | | |
| | RC | | 22·5 | | | 21·0 | 21·0 | 20·0 | 19·4 | 19·5 | 18·7 | 19·8 | 13·0 | 17·7 | | | | |
| TA = 55 | M | | 28·5 | | | 23·6 | 22·1 | 23·2 | 24·1 | 23·8 | 21·4 | 25·9 | 17·0 | 14·7 | | | | |
| | (N) | | (2) | | | (28) | (33) | (31) | (49) | (87) | (15) | (20) | (1) | (3) | | | | |
| | RC | | | | 22·0 | 22·0 | 21·3 | 21·7 | 20·9 | 20·3 | 20·8 | 21·4 | 24·0 | 11·0 | | | | |
| TA = 60 | M | | | | 27·0 | 22·0 | 24·9 | 24·7 | 28·1 | 25·5 | 26·2 | 24·1 | 28·0 | 8·0 | | | | |
| | (N) | | | | (3) | (2) | (20) | (27) | (76) | (39) | (48) | (14) | (1) | (1) | | | | |
| | RC | | | | | | | 24·3 | 25·5 | 22·4 | 20·0 | | | | | | | |
| A = 70 | M | | | | | | | 33·4 | 30·9 | 29·5 | 27·0 | | | | | | | |
| | (N) | | | | | | | (7) | (54) | (42) | (2) | | | | | | | |

(N) = Number in this category.
RC = Reading Comprehension mean score.
M = Mathematics mean score.
TA = Total Ability raw score (0 to 80).

and 75th centiles could qualify. As well as the above global comparison of the D-group with the population, we therefore made comparisons within certain selected total ability scores. Over a wide range of such total ability scores it was found that for both reading comprehension and mathematics average achievement scores were lower at both extremes of V-NV differences within a given total ability. In support of the predominance of higher NV abilities among our D-group, it was also found that for any given total ability score achievement scores generally increased with the V-component to an optimum and then declined fairly sharply. This is shown in Table 5.

## DISCUSSION

Studies which examine underachievement and large V-NV ability differences usually identify underachievers and then measure their ability levels. In this way researchers have often found that certain groups of underachievers (e.g., reading disabled) contain more large V-NV ability differences than would a normal sample, but the number of such cases is usually not overwhelming. However, it has often been remarked that such groups are heterogeneous with regard to the underlying causes of the underachievement.

This study, by contrast, began by identifying children with large V-NV ability differences and then showed that severe underachievement occurred in 22 per cent of the sample, while 61 per cent scored below prediction in both reading comprehension and mathematics. Deficits in verbal ability were found in all severely underachieving boys, while deficits in non-verbal ability were also found in girls severely underachieving in mathematics.

Our results suggest that at least some large V-NV ability differences are not just the natural extremes of a continuum, and that they may have disturbing consequences in terms of school achievement. Ideally, a longitudinal study is now needed to confirm that large V-NV differences and underachievement are correlated through time. It is possible that relative strengths of V and NV abilities fluctuate during development with corresponding variation in levels of achievement in reading and mathematics. In this case it is possible that persistently large V-NV differences are predictive of long-term underachievement and of neurological abnormality.

Although the tests used in our study only tapped a very narrow range of both verbal and non-verbal behaviour, and that at a single point in time, we feel that the size of the population considered and the potential educational importance of our results justify their report. These results suggest that a similar study, based on more comprehensive measures of V and NV abilities, would be worthwhile.

Correspondence and requests for reprints should be addressed to Dr. Joyce E. Whittington, 63 Silver Street, Nailsea, Bristol, BS19 2DS, England.

## REFERENCES

Benson, D. F., and Geschwind, N. (1970). Developmental Gerstmann syndrome. *Neurology*, 20, 293-298.

Berk, R. A. (1982). Verbal-Performance IQ Discrepancy Score: a comment on reliability, abnormality and validity. *J. clin. Psychol.*, 38, 638-641.

Butler, N. R., and Bonham, D. G. (1963). *Perinatal Mortality*. Edinburgh: Livingstone.

Crosby, R. M. N. (1976). *Reading and the Dyslexic Child*. London: Souvenir Press.

Davie, R. (1973). Eleven years of childhood, *Stat. News*, 33, 1-7.

Davie, R., Butler, N. R., and Goldstein, H. (1972). *From Birth to Seven*. London, Longman/National Children's Bureau.

Douglas, J. W. B. (1964). *The Home and the School*. London: MacGibbon and Kee.

Fogelman, K. (1976). *Britain's Sixteen-Year-Olds*. London: National Children's Bureau.

Kinsbourne, M., and Warrington, E. K. (1963a): Developmental factors in reading and writing backwardness. *Br. J. Psychol.*, 54, 145-156.

Kinsbourne, M., and Warrington, E. K. (1963b). The developmental Gerstmann syndrome. *Archs Neurol.*, 8, 490-501.

Money, J. (1964): Two cytogenetic syndromes. Psychologic comparisons: intelligence and specific factor quotients. *J. psychiat. Res.*, 2, 223-231.

RESKE-NIELSEN, E., *et al.* (1982). A neuropathological and neuropsychological study of Turner's syndrome. *Cortex,* 18, 181-190.

RUTTER, M., and YULE, W. (1975). The concept of specific reading retardation, *J. child Psychol. Psychiat.*, 16, 181-198.

THORNDIKE, R. L. (1963). *The Concepts of Over- and Under-Achievement*, New York: Teachers College, Columbia University.

(*Manuscript received* 19*th October,* 1987)

*Br. J. educ. Psychol.*, 58, 212-216, 1988

# AN INVESTIGATION OF THE VALIDITY OF THE BRITISH ABILITY SCALES WITH RESPECT TO THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN-REVISED AND THE WIDE RANGE ACHIEVEMENT TEST-REVISED ON A GROUP OF CANADIAN CHILDREN

BY JOHN COOK

*(Senior Educational Psychologist, West Sussex County Council)*

SUMMARY. One aspect of the validity of the BAS was assessed for 20 children, mean age 10 years, from British Columbia, comparing their BAS scores with those for the WISC-R and WRAT-R. Correlations were found to be high, especially for general ability measures and for verbal IQ scores. Lower but still significant correlations were found between all other scales and the Visual IQ of the BAS, suggesting that the Visual IQ may be more independent of verbal factors than the Performance IQ of the WISC-R. When using the BAS in North America, it is suggested that individual scores will need to be adjusted upwards by about 3 or 4 points.

## INTRODUCTION

The British Ability Scales (Elliott *et al.*, 1983) are now widely used in Great Britain, but, as Ward and Outhred (1986) note, there is little evidence of any marked international interest. However, in the current Mental Measurements Year Book (Wright and Stone, 1985) they receive a recommendation for more general use by two North American reviewers, albeit with a strong rider about the relative lack of sufficient reliability or validity data in one case. The Scales are also now recorded in a North American publishers' catalogue and so are available on that continent, being developed and standardised for US use as the Differential Ability Scales.

On a recent exchange to British Columbia as an educational psychologist sponsored by the League for the Exchange of Commonwealth Teachers, the author decided to examine the use of the BAS in a Canadian setting and in addition add to the sparse validity data. Working within a School District Service has advantages and disadvantages for such a study. Whilst children are readily available as part of the regular assessment programme required by schools, there is little possibility of obtaining a random sample of the population. Children referred are likely to have learning problems of some kind, giving scores skewed towards the lower part of the ability or achievement range.

In addition, therefore, to administering the BAS on the group of Canadian children, it was decided to incorporate into the study a parallel ability scale of long standing, the WISC-R (Wechsler, 1974). This would add depth to the study by providing a base line for the group assessed and hence an inter-test validity measure.

## METHOD

### *Sample*

There were 20 children in the group assessed. All had been referred to the Special Services Department of School District 35, Langley, BC, Canada as needing assessment by an educational psychologist. The primary reason for referral as stated by school in 10 cases was general problems with learning, in six was to do with behaviour, and in four with both learning and behaviour components. All were boys with a mean age of 10·1 years, a range from 8 years to 13·2 years and a standard deviation of 1·7 years.

No attempt was made to select the sample except that they had to be of at least 8 years of age. They were children for whom each of the 10 schools in one geographic area of the School District had highest priority for assessment by an educational psychologist. All were seen as part of the regular service which the Special Services Department provided for the schools.

### *Procedure*

All children were given both the BAS and the WISC-R tests within their regular school. Order of presentation was varied randomly with an equal frequency of each test having the first presentation. A break from the interview of ½/¾ hr. was provided between each test administration for every child.

No attempt was made to score the BAS protocols until after all 20 pairs of presentations had been completed. Presentation of results to parents and schools meant that the WISC-R had to be scored and a report written on each child.

Ten subtests of the WISC-R were administered: Information, Similarities, Arithmetic, Vocabulary and Digit Span of the Verbal Scale (omitting Comprehension) and Picture Completion, Picture Arrangement, Block Design, Object Assembly and Coding on the Performance Scale. The Comprehension subscale was omitted from the WISC-R because it is a subtest which the author finds most difficult to score objectively. Seven scales of the BAS were used; Similarities, Recall of Digits and Word Definitions, from which a Verbal IQ was derived; Speed of Information Processing, Matrices and Block Design-Level produced a Visual IQ. Word Reading was also given.

Two adaptations of the BAS were made for Canadian children, both substitutes in the Similarities Scale. "Pants" was substituted for "trousers" and "sardines" for "pilchards". Otherwise instructions and scoring procedures were followed exactly.

In addition to the tests administered as part of the study, a score was also available for each child for the reading scale of the Wide Range Achievement Test-Revised (Jastak and Wilkinson, 1986) usually given by the school's learning assistance teacher. All WRAT-R scores were obtained within one month of the date of the educational psychologist's interview.

## RESULTS

A comparison between the BAS and the WISC-R IQ's showed the WISC-R to score about 3 or 4 points higher than the corresponding score for the BAS. This difference was significant ($P < 0{\cdot}05$) for the general ability and verbal scores, but only approached significance ($P < 0{\cdot}2$) when the non-verbal measures were compared.

The Pearson product-moment correlations between the various IQ scores were mostly impressively high ($P < 0{\cdot}01$) save for those comparisons with the Visual IQ of the BAS. Although the BAS Visual IQ correlations remain significant, the level of confidence ($P < 0{\cdot}05$) is generally lower (See Table 2). Whilst there is close conformity between within-test correlations found for the WISC-R by Wechsler (1974), no such data are presently published for within-test comparisons for the BAS. A validation study reported by Elliott (1983) between the Short Form IQ of the BAS and the WISC-R found intercorrelations rather lower than those found here, but a similar correlation between the BAS General IQ and the WISC-R Performance IQ.

TABLE 1

Means and Standard Deviations of Quotients for BAS, WISC-R and WRAT-R

| | BAS | WISC-R | Difference between means (two-tailed) |
|---|---|---|---|
| General/ Full Scale | 94·6 (12·7) | 98·1 (13·9) | P < 0·05 |
| Verbal/ Verbal Scale | 92·0 (17·5) | 96·4 (15) | P < 0·025 |
| Visual/ Performance Scale | 97·6 (11·4) | 100·4 (11·4) | P < 0·2 |

| BAS Word Reading | WRAT-R Reading | Difference between means (two-tailed) |
|---|---|---|
| 95·7 (17·4) | 92·3 (18·4) | P < 0·05 |

TABLE 2

INTERCORRELATIONS BETWEEN WISC-R AND BAS IQ SCORES

| | | WISC-R | | | BAS | | |
|---|---|---|---|---|---|---|---|
| | | Full Scale | Verbal Scale | Perf. Scale | General | Verbal | Visual |
| WISC-R | Full Scale IQ | — | | | | | |
| | Verbal Scale IQ | 0·96** | — | | | | |
| | Performance Scale IQ | 0·93** | 0·81** | — | | | |
| BAS | General IQ | 0·91** | 0·91** | 0·74** | | | |
| | Verbal IQ | 0·87** | 0·92** | 0·71** | 0·94** | — | |
| | Visual IQ | 0·57* | 0·48* | 0·47* | 0·71** | 0·45* | — |

*P < 0·05 **P < 0·01 (one-tailed)

TABLE 3

MEANS AND STANDARD DEVIATIONS OF T-SCORES (M = 50; SD = 10) FOR SUBSCALES OF THE BAS

| | Scale | Mean | (SD) |
|---|---|---|---|
| Verbal Scales | Similarities | 45·8 | (9·1) |
| | Recall of Digits | 40·5 | (16·5) |
| | Word Definitions | 46·0 | (13·4) |
| Visual Scales | Speed of Information Processing | 53·3 | (9·8) |
| | Matrices | 45·2 | (12·4) |
| | Block Design-level | 46·6 | (7·7) |

TABLE 4

MEANS AND STANDARD DEVIATIONS OF SUBTESTS OF THE WISC-R

| | Subtest | Mean | (SD) |
|---|---|---|---|
| Verbal | Information | 8·5 | (2·4) |
| | Similarities | 10·2 | (3·7) |
| | Arithmetic | 9·3 | (2·8) |
| | Vocabulary | 9·3 | (3·9) |
| | Digit Span | 8·4 | (2·3) |
| Performance | Picture Completion | 9·8 | (3·9) |
| | Picture Arrangement | 10·1 | (3·6) |
| | Block Design | 9·6 | (3·3) |
| | Object Assembly | 9·7 | (2·3) |
| | Coding | 8·6 | (2·3) |

No clear profile emerges from the mean scores of the subscales of the BAS (Table 3) and, although there is some suggestion of a reading disability like profile (Sattler, 1982) for the WISC-R (Table 4), the sample cannot be characterised as reading disabled. The highest subscale score for the group on the BAS was Speed of Information Processing, at which children with a specific reading disability tend to perform poorly (Thomson, 1982).

Often important clinically are differences between verbal and non-verbal scores. The comparisons for each individual in the sample between their BAS Verbal/Visual IQs and between their WISC-R Verbal/Performance IQs showed that 17 of the 20 children had IQ differences of less than 12 points. The remaining three, all of whom were referred for emotional or behavioural difficulties, differed by more than 25 points. Caution, therefore, needs to be exercised when making interpretations about individual verbal/non-verbal differences when emotional or behavioural factors are indicated at referral.

The BAS Word Reading scores correlated highly with the WRAT-R standard scores for reading ($r = 0{\cdot}89$, $P < 0{\cdot}01$) but the means differed significantly ($P < 0{\cdot}05$). Incorporating the reading scores with the respective IQ scores shows an apparent contradiction. Whilst for the BAS IQs and Word Reading scores there is no difference, for the WISC-R Full Scale IQs and the WRAT-R scores the two means differ significantly ($P < 0{\cdot}01$). However, the BAS IQ and reading scores were standardised on the same sample and can therefore be directly compared whereas the WISC-R and the WRAT-R were not.

## DISCUSSION

The results of this study show the BAS to be a valid instrument paralleling the scores of the WISC-R for a British Columbian population. Whilst 20 subjects in one part of the North American continent is too small a sample to enable major generalisations to be made, the magnitudes of the correlations are impressive, suggesting that the BAS is likely to be a valid instrument for use more widely with children. It should be remembered, however, that the group was small, composed solely of boys and was a referred population rather than randomly selected. Nevertheless, there is good evidence to support the contention of Wright and Stone (1985) that the BAS is likely to be effective in the US and Canada.

The 3- to 4-point difference between scores, although statistically significant, is not sufficient to affect the use of the BAS provided an adjustment of that amount is made to the score.

A rider needs to be added about verbal/non-verbal differences. The lower correlations of the BAS Visual IQ with all other IQs suggests that this may be a better guide to the spatial factor, and more independent of verbal skills than the Performance IQ of the WISC-R. Hence it may be a more valid guide in evaluating verbal/non-verbal differences for an individual than the Verbal/Performance IQ difference of the WISC-R. The ability of the BAS go to beyond this difference, to investigate the reality of differences between individual subscales, puts the investigation of within-child differences on a much sounder statistical basis than the speculation about subtest differences for the WISC-R often seen in clinical reports on individual children. The addition of the standard error of measurement throughout the normative tables of the BAS extends the practical robustness of the procedure.

The highest subscale score obtained for the BAS was Speed of Information Processing which was significantly higher than all other subscales and the only one to score above the mean of the British normative population. As a number cancelling task, where speed and accuracy are important, it may be that the North American motivation, self-awareness and will to succeed is a significant cultural factor which particularly affects this subscale. A small-scale study has been instigated this further (Funk and Cook, in preparation).

ACKNOWLEDGMENTS I should like to thank Dr Shirley McBride and Mr Arnold Funk of Langley School District, Mr Keith Pickup of West Sussex Local Education Authority, and the League for the Exchange of Commonwealth Teachers for their assistance in the completion of this project. I should also like to acknowledge the interest and assistance of the anonymous referees who guided me closer to publishable rigour.

Correspondence and requests for reprints should be addressed to Mr John Cook, Education Department, West Sussex County Council School Psychological Service, 72-74 Wick Street, Littlehampton, West Sussex, BH17 7JS, England.

## REFERENCES

ELLIOTT, C. D. (1983). *British Ability Scales, Revised Edition. Technical Handbook.* Windsor: NFER-Nelson.

ELLIOTT, C. D., MURRAY, D. J., and PEARSON, L. S. (1983). *The British Ability Scales, Revised Edition.* Windsor: NFER-Nelson.

EMBRETSON, S. (1985), Review of the British Ability Scales. In MITCHELL, J. V. (Ed.), *The Ninth Mental Measurements Yearbook.* Lincoln: University of Nebraska Press.

FUNK, A., and COOK, J. Differences between scores for Canadian and British children for the Speed of Information Processing Scale of the BAS (in preparation).

JASTAK, S., and WILKINSON, G. (1986). *The Wide Range Achievement Test — Revised.* Windsor: NFER-Nelson.

SATTLER, J. M. (1982). *Assessment of Children's Intelligence and Special Abilities.* Boston: Allyn and Bacon.

TEXAS INSTRUMENTS (1983). *Owners Manual, TI-55111.* Bedford: Texas Instruments.

THOMSON, M. E. (1982). The assessment of children with specific reading difficulties (dyslexia) using the British Ability Scales. *Br. J. Psychol.*, 73, 461-478.

WARD, J., and OUTHRED, L. (1986). The use of the British Ability Scales with a group of Australian school children — some preliminary data. *Br. J. educ. Psychol.*, 56, 84-87.

WECHSLER, D. (1974). *The Wechsler Intelligence Scale for Children — Revised.* New York: The Psychological Corporation.

WRIGHT, B. D., and STONE, M. A. (1985). Review of British Ability Scales. In MITCHELL, J. V. (Ed.), *The Ninth Mental Measurements Yearbook.* Lincoln: University of Nebraska Press.

(*Manuscript received* 28*th October,* 1986)

*Br. J. educ. Psychol.*, 58, 217-222, 1988

# PATTERNS OF PERFORMANCE ON THE BRITISH ABILITY SCALES FOR A GROUP OF CHILDREN WITH SEVERE READING DIFFICULTIES

BY M. F. McKAY
*(Institute of Catholic Education, Victoria, Australia)*
AND M. D. NEALE
*(Victoria, Australia)*

SUMMARY. Patterns of subtest performance on the British Ability Scales (Elliott *et al.*, 1979) for a group of 24 children attending a reading disability clinic were analysed in the light of the report by Thomson (1982). The present findings generally support those of the earlier study, and add further credence to the more general attempts of diagnosticians to understand the multi-faceted nature of reading problems.

## INTRODUCTION

A large part of the clinical and empirical research on the Wechsler Intelligence Scale for Children — WISC (Weschsler, 1949, 1974) has concerned itself with the psychoeducational utility of using the WISC in the diagnostic process. Characteristic profiles of reading disabled children have received particular attention (Belmont and Birch, 1966; Huelsman, 1970; Ackerman *et al.,* 1971; Bannatyne, 1968, 1971; Rugel 1974a, 1974b; Vance *et al.,* 1978; Thomson and Grant, 1979; Moore and Wielan, 1981; McKay *et al.,* 1985 among others).

More recently the British Ability Scales (BAS) (Elliott *et al.*, 1979) have been used to examine the pattern of strengths and weaknesses in cognitive processes displayed by children with specific reading problems (Thomson, 1982). Amongst other things, this study provided findings which suggested that such children across a variety of age groups have deficity (significantly lower T-scores) in Speed of Information Processing, Word Reading, Basic Arithmetic, Immediate and Delayed Recall and Recall of Digits as measured by the BAS. Thomson also reported that the results appeared to be related to a specific learning problem rather than overall low ability and that there was no evidence of "compensatory" higher scores in "non-deficit" areas, unlike some of the findings with the WISC (Bannatyne, 1971; Thomson and Grant, 1979). The author then went on to relate the diagnostic implications of such deficits to previous supporting evidence from both clinical and empirical studies and to discuss the use of the expectancy tables to identify discrepancies between "potential ability" and reading attainment.

This paper aims to report on the BAS profiles of a group of children with specific reading difficulties who were attending a special remedial programme at the Krongold Centre, Monash University. Confirmation of Thomson's analysis concerning patterns of performance on the BAS for children with severe reading difficulties is its primary concern.

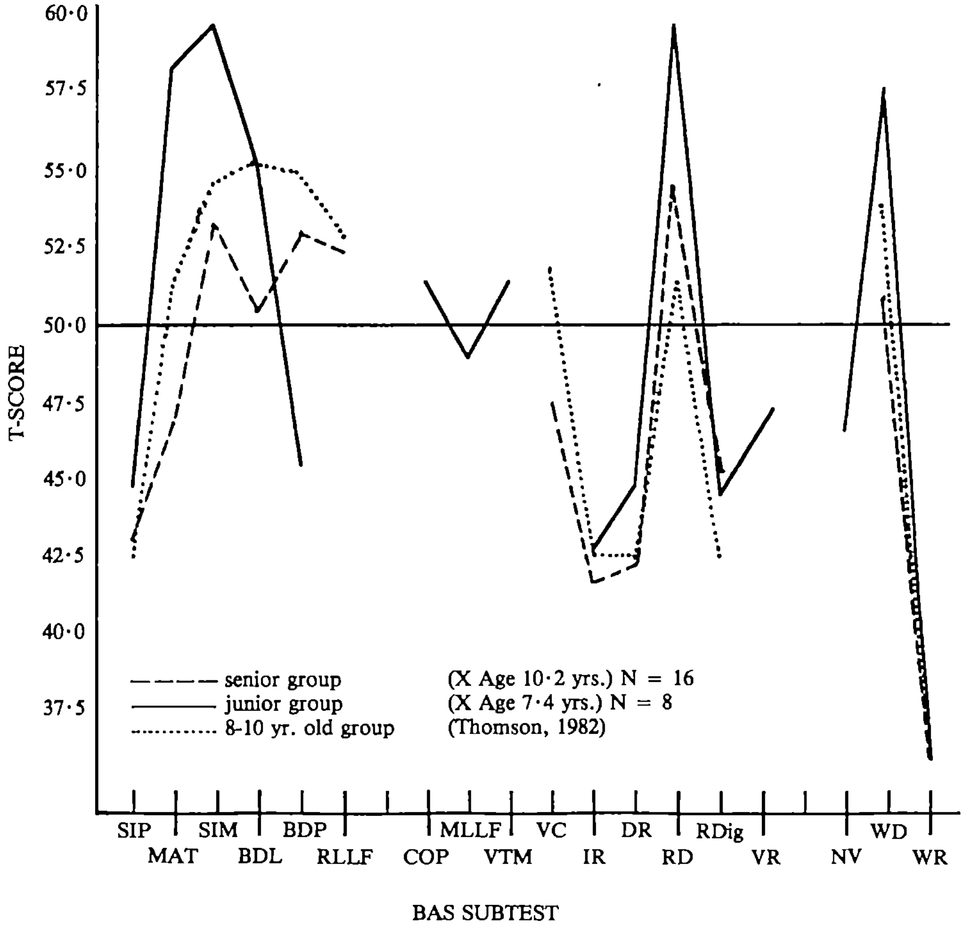
## METHOD

Twenty-four children who had displayed persistent and chronic difficulties with reading and other language skills constituted the sample for this study. They were drawn from a larger pool of children who attended the Krongold Centre for Exceptional Children at Monash University throughout 1985, for assistance with their learning problems. Over the past 10 years this Centre has provided services through remediation and enrichment programmes and conducted research in various fields, e.g., developmental disabilities, giftedness, literacy skills and reading difficulties. The ages of the children ranged from 6 years 11 months to 11 years 11 months. All were participating in two separate programmes of remediation conducted sessionally at the Centre, one designed for the more junior members of the group as a "fresh start" programme (N = 8) and the other for the older members (N = 16). The mean age of the junior group was 7 years 4 months (SD = 6 months) and that of the senior group 10 years 2 months (SD = 15 months). All children displayed the general features of specific reading retardation (Rutter, 1978) including directional confusion or trouble in sequencing, late language development, left/right confusion, writing and spelling difficulties and, more often than not, family histories of similar difficulties.

The children had been screened for visual impairments, emotional disorders, English as a second language problem, general health difficulties and obvious signs of brain damage. All

FIGURE 1

Profiles of Mean Subtest Performance for Junior and Senior Groups and 8-10-Year-Old Group (Thomson, 1982)



KEY

SIP = Speed of Information Processing
MAT = Matrices
SIM = Similarities
BDL = Block Design Level
BDP = Block Design Power
RLLF = Rotation of Letter-Like Forms
COP = Copying
MLLF = Matching Letter-Like Forms
VTM = Verbal-Tactile Matching

VC = Visualisation of Cubes
IR = Immediate Visual Recall
DR = Delayed Visual Recall
RD = Recall of Designs
RDig = Recall of Digits
VR = Visual Recognition
NV = Naming Vocabulary
WD = Word Definitions
WR = Word Recognition

children in the older group were at least 18 months retarded from chronological age in Reading Accuracy using the Neale Analysis of Reading Ability — Revised. Seven out of eight of the junior group were not reading, i.e., could not obtain a basal score on the same test, which requires at least 11 words to be read accurately from a passage 26 words in length. Mean IQ was in the normal range for both age groups, with the junior group scoring 101·5 on the Full Scale of the BAS and 102 and 99·1 for the Visual and Verbal Scales respectively. The senior group obtained a mean Full Score BAS IQ of 103·8 with results of 102·5 and 102·8 for Visual and Verbal IQ Scales respectively.

*Procedure*

All children were administered the age-appropriate BAS Scales and the Neale Analysis of Reading Ability — Revised, within two weeks of entry to their remedial programme. A number of criterion-referenced tests and tests of academic achievement were also administered at the same time. Full details of the age appropriate subtests of the BAS are provided in the Manual (Elliott *et al.*, 1983). The Neale Analysis of Reading Ability — Revised (Neale, 1986) is a recent version of the well-known oral reading test which consists of three parallel forms each of six graded pasaages of prose forming a continuous reading scale for children aged from 6 to 12+ years. The test measures Rate, Accuracy and Comprehension of oral reading. Local Australian norms were used.

## RESULTS AND DISCUSSION

Figure 1 depicts the T-score profiles for the two age groups, based on a mean T-score of 50 and a standard deviation of 10 for all scales. The profile of performance for Thomson's 8-10-year-old group (Thomson, 1982) is superimposed over these results.

It can be observed that both groups of children appeared to do poorly in the subtests Speed of Information Processing, Immediate Recall, Delayed Recall, Recall of Digits and in Word Reading. These results are directly in line with the findings of Thomson (1982) with the exception of Basic Arithmetic. This subtest was not administered to our sample. Detailed means and standard deviations for individual subtest T-score performances, for the junior group and senior group, are separately reported in Table 1.

TABLE 1

MEAN BAS T-SCORE PERFORMANCES FOR TWO GROUPS OF POOR READERS, JUNIORS AND SENIORS

| | *Juniors N = 8* | | *Seniors N = 16* | |
|---|---|---|---|---|
| *BAS Sub-test* | *Mean* | *SD* | *Mean* | *SD* |
| ↓ SIP | 44·5 | (14·8) | 43·6 | (9·8) |
| ↑ MAT | 57·8 | (4·8) | 47·0 | (8·8) |
| ↑ SIM | 59·3 | (5·3) | 53·2 | (11·8) |
| ↑ BDL | 55·5 | (6·0) | 50·4 | (12·6) |
| BDP | 45·8 | (14·9) | 52·2 | (13·5) |
| RLLF | — | — | 51·6 | (9·8) |
| COP | 51·5 | (9·5) | — | — |
| MLLF | 49·1 | (4·4) | — | — |
| VTM | 51·3 | (12·9) | — | — |
| VC | — | — | 47·0 | (11·7) |
| ↓ IR | 42·8 | (14·4) | 42·1 | (9·3) |
| ↓ DR | 44·5 | (11·6) | 42·3 | (9·4) |
| ↑ RD | 59·6 | (11·0) | 54·2 | (9·3) |
| ↓ RDIG | 44·0 | (11·6) | 45·3 | (12·4) |
| VR | 47·0 | (10·9) | — | — |
| NV | 46·5 | (6·4) | — | — |
| ↑ WD | 58·6 | (7·8) | 50·5 | (8·8) |
| ↓ WR | 36·3 | (7·4) | 36·6 | (6·1) |

↓ Indicates sub-tests on which *both* Junior and Senior groups were below the norm.

↑ Indicates sub-tests on which *both* Junior and Senior groups were at or above the norm.

A two-way repeated measures ANOVA revealed significant within-group differences between subtests $F(10,231) = 7{\cdot}3$ $P < 0{\cdot}001$. No significant overall effects however, were found between the junior and senior groups $F(1,10)$ $F = 3{\cdot}5$. Accordingly, results for the junior and senior groups were combined, in order to explore the pattern of differences between subtests. Those subtests which were at or above the norm on the one hand, or those which were below the norm, have been separately indicated in Table 1 above (after Thomson, 1982).

The significance of the differences between these particular subtest means was then assessed (setting $\alpha$ at 0·05 level) using the age appropriate tables for comparisons of discrepancies between T-scores (Elliott *et al.*, 1983). Of those subtest mean T-scores which were low, a number proved to be significantly so (See Table 2).

Word Reading was significantly weaker than all other subtests below the norm as would be expected with Speed of Information Processing being also significantly lower than most other subtests at or above the norm. The exceptions to this latter result were mainly other

TABLE 2

OUTCOMES OF CONTRASTS BETWEEN OBTAINED MEAN T-SCORE AND "EXPECTED DISCREPANCIES" BETWEEN VARIOUS MEAN T-SCORES (AFTER ELLIOTT *et al.*, 1983)

| *Contrast* | **Expected Diff.* | *Actual Diff* | *P* |
|---|---|---|---|
| SIP vs SIM | 11 | 11 | <0·05 |
| SIP vs BDL | 10 | 8 | NS |
| SIP vs RD | 10 | 12 | <0·05 |
| SIP vs WD | 9 | 9 | <0·05 |
| SIP vs MAT | 9 | 6 | NS |
| SIP vs BDP | 11 | 7 | NS |
| IR vs SIM | 17 | 11 | NS |
| IR vs BDL | 17 | 10 | NS |
| IR vs RD | 16 | 13 | NS |
| IR vs WD | 16 | 16 | <0·05 |
| IR vs MAT | 16 | 8 | NS |
| IR vs BDP | 18 | 8 | NS |
| DR vs SIM | 17 | 12 | NS |
| DR vs BDL | 16 | 9 | NS |
| DR vs RD | 16 | 13 | NS |
| DR vs WD | 14 | 10 | NS |
| DR vs MAT | 16 | 7 | NS |
| DR vs BDP | 17 | 8 | NS |
| RDig vs SIM | 12 | 10 | NS |
| RDig vs BDL | 12 | 7 | NS |
| RDig vs RD | 11 | 11 | <0·05 |
| RDig vs WD | 9 | 8 | NS |
| RDig vs MAT | 11 | 5 | NS |
| RDig vs BDP | 13 | 6 | NS |
| WR vs SIM | 10 | 18 | <0·05 |
| WR vs BDL | 9 | 15 | <0·05 |
| WR vs RD | 8 | 19 | <0·05 |
| WR vs WD | 8 | 16 | <0·05 |
| WR vs MAT | 7 | 14 | <0·05 |
| WR vs BDP | 10 | 14 | <0·05 |

timed subtests, Block Design Level and Power. The lack of a significant difference between Speed of Information Processing and Matrices, although not a timed subtest, is interesting since spatial synthesis and/or problem solving involving the selection of an appropriate strategy based on sequential scanning is involved in both subtests. Of the remaining subtests below the norm, only Immediate Visual Recall and Recall of Digits were significantly different from any of the subtest means above the norm. In the first instance the difference between Immediate Visual Recall (naming objects recalled from a picture card presentation) and Word Definitions would seem an unexpected outcome since both involve semantic associations and to some extent labelling. The essential difference between the two subtests appears to be the memory component involved in Immediate Visual Recall. This has often been characterised as reflecting a deficit in attention, concentration or "immediate memory" (Pincus and Glaser, 1966; Kaufman, 1981; McKay *et al.*, 1985). In regard to the difference between Recall of Digits (lower mean T-score) and Recall of Designs (higher mean T-score), the auditory component involved in either the initial input or the recall marks the difference in the memory component of each task. The poorer performance on the Recall of Digits task is not an unexpected outcome, since many studies have reported similar results (e.g., Egeland *et al.*, 1970; Naidoo, 1972; Fuller and Friedrich, 1974; Thomson and Wilsher, 1978).

Overall the subtests on which this group performed at least at an average level, i.e., Block Design (Power and Level), Recall of Designs and Matrices, would be in line with findings from studies using the WISC-R for similar tasks (e.g., Bannatyne, 1971; Rugel, 1974a; McKay *et al.*, 1985). These studies have found that the ability to manipulate objects directly or symbolically in multi-dimensional space is typically the strongest area of performance for these children, suggesting therefore that the problem they have is related to tasks for which the left cerebral hemisphere has "executive" control. These findings are also in line with those of Thomson (1982).

In summary, these findings support the BAS profile of abilities for children with specific reading difficulties as reported by Thomson. Such intra-individual differences of abilities for groups of LD children should be acknowledged and serve to remind us that good intelligence can remain intact in spite of severe disability in one or more specific cognitive dimensions. The instrument, while yet to achieve the currency of other tests such as the WISC, appears to be a valuable additional tool for the psychologist concerned with differential diagnosis relating to reading difficulties.

Correspondence and requests for reprints should be addressed to Dr. M. D. Neale, P.O. Box 141, Elwood, Victoria 3184, Australia.

## REFERENCES

ACKERMAN, P. T., PETERS, J. E., and DYKMAN, R. A. (1971). Children with specific learning difficulties: WISC profiles. *J. learn. Disabil.*, 4, 150-166.

BANNATYNE, A. (1968). Diagnosing learning disabilities and writing remedial prescriptions. *J. learn. Disabil.*, 1, 28-35.

BANNATYNE, A. (1971). *Language, Reading and Learning Disabilities*. Springfield, Ill.: Charles C. Thomas.

BELMONT, L., and BIRCH, H. G. (1966). The intellectual profile of retarded readers. *Percept. mot. Skills*, 22, 787-816.

EGELAND, B., DI NELLO, M., and CARR, D. (1970). The relationship of intelligence, visual-motor, psycholinguistic and reading readiness skills with achievement. *J. educ. psychol. Measurement*, 30, 451-458.

ELLIOTT, C. D., MURRAY, D. J., and PEARSON, L. S. (1979, 1983). *The British Ability Scales* (Revised 1983). Slough: NFER.

FULLER, G. B., and FRIEDRICH, D. (1974). Three diagnostic patterns of reading disabilities. *Academic Therapy*, 10 (2), 219-231.

HUELSMAN, C. B. (1970). The WISC subtest syndrome for disabled readers. *Percept. mot. Skills*, 30, 535-550.

KAUFMAN, A. S. (1981). The WISC-R and learning disabilities assessment: state of the art. *J. learn. Disabil.*, 14, 520-526.

MOORE, D. W., and WIELAN, O. P. (1981). WISC-R scatter indices of children referred for reading diagnosis. *J. learn. Disabil.*, 14, 511-514.

MCKAY, M. F., NEALE, M. D., and THOMPSON, G. B. (1985). The predictive validity of Bannatyne's WISC categories for later reading achievement. *Br. J. educ. Psychol.*, 55, 280-287.

NAIDOO, S. (1972). *Specific Dyslexia*. London: Invalid Children's Aid Association/Pitman.

NEALE, M. D. (1986). *The Neale Analysis of Reading Ability — Revised.* Melbourne: ACER.

PINCUS, J. H., and GLASER, G. H. (1966). The syndrome of "minimal brain dysfunction in childhood". *New England J. Medicine*, 275, 27-35.

RUGEL, R. P. (1974a). WISC subtest scores of disabled readers: review with respect to Bannatyne's recategorisation. *J. learn. Disabil.*, 7, 48-55.

RUGEL, R. P. (1974b). The factor structure of the WISC in two populations of disabled readers. *J. learn. Disabil.*, 7, 581-585.

RUTTER, M. (1978). Prevalence and types of dyslexia. In BENTON, A. C., and PEARL, D. (Eds.), *Dyslexia: An Appraisal of Current Knowledge.* New York: Oxford University Press.

THOMSON, M. E. (1982). The assessment of children with specific reading difficulties (dyslexia) using the British Ability Scales. *Br. J. Psychol.*, 73, 461-478.

THOMSON, M. E., and GRANT, S. (1979). The WISC subtest profile of the dyslexic child. In NEWTON, M. J., THOMSON, M. E., and RICHARDS, I. (Eds.), *Readings in Dyslexia.* Wisbech, Cambridgeshire: Learning Development Aids.

THOMSON, M. E., and WILSHER, C. (1978). Some aspects of memory in dyslexics and controls. In GRUNEBERG, M. M., MORRIS, P. E., and SYKES, R. N. (Eds.), *Practical Aspects of Memory.* London: Academic Press.

VANCE, H. B., WALLBROWN, F. H., and BLAKE, J. (1978). Determining WISC-R profiles for reading disabled children. *J. learn. Disabil.*, 11, 657-661.

WECHSLER, D. (1949). *The Wechsler Intelligence Scale for Children.* New York: The Psychological Corporation.

WECHSLER, D. (1974). *Wechsler Intelligence Scale for Children — Revised.* New York: The Psychological Corporation.

(*Manuscript received* 8*th December* 1986)

# CONFERENCE ABSTRACTS

*Br. J. educ. Psychol.*, 58, 223-230, 1988

Papers delivered at the Annual Conference of the Education Section of the British Psychological Society, held in the St Nicholas Hotel, Scarborough, October 1987.

ALLEN, M., CCDU, University of Leeds.

## EXPERIENTIAL METHODS IN EDUCATION — PURPOSES, POSSIBILITIES AND PROBLEMS

This session is intended as a discussion of some issues arising from the increasingly popular use of experiential methods in education. The session will open with a brief input by P. Tomlinson and M. Allen and then open discussion in which it is hoped that people will participate fully.

FAWCETT, A., Department of Psychology, University of Sheffield.

## DEVELOPMENTAL DYSLEXIA — HOW FAR DOES READING REMEDIATION EXTEND?

Early identification enables some children to overcome their reading problems. However, memory (Jorm, 1983) and organisational (Torgeson and Houck, 1980) difficulties are also reported. Is remediation specific to reading or affected by changes in general abilities? A temporal sequencing task, ordering the months, normally acquired at age 6-7 (Friedman, 1986) remains a problem for dyslexics (Miles, 1982), The performance of remediated dyslexics was compared with chronological age controls and other dyslexics.

*Experiment 1*

Reciting the months forwards and backwards. The remediated dyslexics ($N = 10$, CA 10 yrs 5 m, IQ 120) were significantly worse than

(a) CA controls ($N = 8$, CA 10 yrs 8 m, IQ 108 $U = 10$ $P < 0{\cdot}01$)
(b) younger controls ($N = 8$, CA 8 yrs 5 m, IQ 106) $U = 14$ $P < 0{\cdot}05$)

Reciting the months backwards was the main difficulty.

(c) CA controls $U = 12$ $P < 0{\cdot}05$
(d) younger controls $U = 14{\cdot}5$ $P < 0{\cdot}05$

Remediated dyslexics have attained stage 1, the verbal ordering of the months, but fail at stage 2, the spatial manipulation of the sequence (Friedman, 1986).

However, dyslexia has also been associated with naming difficulties (Denckla and Rudel, 1976). Is the problem one of recall or recognition?

*Experiment 2*

A random visual and auditory presentation of the months. Remediated dyslexics ($N = 8$, CA 11 yrs 4 m, IQ 119) were significantly worse than the control group ($N = 8$, CA 11 yrs 11 m, IQ 108) on

(a) written months backwards $U = 6$ $P < 0{\cdot}01$
(b) spoken months forwards $U = 10$ $P < 0{\cdot}05$
(c) spoken months backwards $U = 12$ $P < 0{\cdot}05$

No significant differences were apparent between remediated and other dyslexics ($N = 9$, CA 12 yrs 4 m, RA 9 yrs 4 m).

Conclusion: Remediated dyslexics show a deficit in temporal ordering, related to sequencing rather than naming difficulties.

GALLOWAY, D., Department of Educational Studies, University of Lancaster.

## BEHAVIOUR PROBLEMS IN SCHOOLS: WHICH TEACHER'S RESPONSIBILITY?

This paper reviews aspects of school organisation, the curriculum and classroom teaching processes which contribute to disruptive behaviour. It is argued that responsibility for pupils with special needs is too often divided between pastoral care staff and the special needs or learning support team, and that this creates an arbitrary split which diverts attention from the school or classroom variables associated with disruptive behaviour. The implications for school and classroom management are discussed with particular reference to the government's proposals for appraisal and for a national curriculum.

HALL, E., and WOOSTER, A., School of Education, University of Nottingham.

## EXPERIENTIAL LEARNING IN TEACHER EDUCATION

Experiential learning is increasingly playing a part in teacher education. This paper reports the outcomes of experienced teachers applying the learning from courses which have been run entirely on experiential lines. These courses have involved structured exercises relating to communication skills, the development of the self-concept and working in groups. Fantasy and drawing exercises are used. Learning is practised and developed in an unstructured small group run on Rogerian lines. Students take a high level of responsibility for their learning in these situations, drawing their own conclusions, identifying and working towards their own goals. In the work setting, the teachers report improved relationships with both students and staff, they become more interested in co-operative learning and move towards a humanistic rather than a custodial attitude to teaching. They also report lower levels of stress in both their personal and working lives. These changes are attributed to the experiential learning process. When the experiential sessions contribute to an award bearing course, academic standards are raised and career prospects are improved.

HONESS, T., Department of Psychology, University College, Cardiff; EDWARDS, A., St. Martin's College, Lancaster.

## MAKING SENSE OF LEAVING SCHOOLING: THE IDENTITY CONSTRUCTIONS OF GIRL SCHOOL LEAVERS

The identity constructions observed in a sample of 160 school leavers supports Harre's notion of the "person as theorist" located within the affordances of specific contexts (Harre, 1983). Case study data from 12 girls drawn from the larger sample suggest that within the locations of school, with girl friends, with boyfriends and with the family, their styles of identity negotiation differ in varying degrees from those of boys from the same communities. Areas in which the girls held theories of personal effectiveness are made evident.

### REFERENCE

HARRE, R. (1983). *Personal Being.* Oxford: Blackwell.

HORNBY, G., Hester Adrian Research Centre, University of Manchester.

## TEACHING BASIC COUNSELLING SKILLS TO TEACHERS

The value of teachers using appropriate basic counselling skills in fulfilling their guidance role with children has long been accepted. More recently, with increasing parental involvement in education, particularly for parents of children with special needs, the importance of possessing these counselling skills has been emphasised.

In order to meet the needs of teachers learning basic counselling skills, a 10-hour practical training course has been developed. The course was designed to be taught in five two-hour sessions, held at weekly intervals. It introduces a simple three-stage model of the counselling process: listening; understanding; and, problem solving. The skills associated with each stage are taught using a combination of role-play, modelling, mini-lectures and discussion.

The course has so far been used with several groups of teachers in training, with practising teachers, other professionals and parents. No formal evaluation of the effectiveness of the course has yet been carried out, but feedback from participants has been overwhelmingly positive.

The paper will present the rationale for, and an overview of the course. Methods of evaluation will be discussed and copies of the course leaders' manual made available.

---

KELLY, M., and RONALDSON, J., Department of Educational Psychology, University of Newcastle.

## EMOTIONAL AND BEHAVIOURAL PROBLEMS IN SCHOOLS: THE NEWCASTLE EXPERIENCE

A course was set up for teachers in mainstream schools in Newcastle who are engaged with children with behavioural difficulties as part of their everyday work. The materials comprised: (a) The Coventry course "Behaviour Difficulties in the Primary School" modified and adapted for Newcastle schools. (b) Additional materials developed by Newcastle psychologists and teachers.

The course was piloted in two ways:

(1) with the entire staff of Kenton Bar Primary School and representatives from the senior management and special needs staff of Kenton Comprehensive to which it feeds;
(2) with the head teachers and representatives of pastoral care/special needs staff of Walker Comprehensive School.

The evaluation, undertaken by Newcastle's SNAP Co-ordinator (M.M.K.) was largely qualitative and involved the use of observations and questionnaires completed by course members. Evaluation results show an extremely positive response by course participants to materials, content, presentations and workshop activities. The measure of its success is mirrored in the demand by each of the participating schools in the Walker pyramid for further courses for those teachers not involved in the pilot study.

Further courses will be set up by means of a consultative model, reflecting the evaluation findings which indicated extreme pressure of time. Such courses will offer a wide menu which will include the original course materials, from which schools may choose according to their own needs.

---

KITZINGER, C., Centre for Social and Moral Education, University of Leicester.

## "IT'S NOT FAIR ON GIRLS": YOUNG WOMEN'S ACCOUNTS OF UNFAIRNESS IN SCHOOL

As part of a larger study of secondary school students' perceptions of unfairness in school, over one thousand girls from a range of different schools provided written accounts of perceived unfairness. Additionally, 80 were interviewed. Compared with boys, girls' complaints of injustice are more frequently body-centred — relating to sexual harassment, school uniform regulations and dress codes, compulsory showers and sports, and restricted access to lavatories. These results are discussed in the context of the social meanings attached to young women's bodies, as deviant, as liabilities and as commodities.

LINDSAY, G., Psychological Service, Sheffield LEA, and Division of Education, University of Sheffield.

## APPLYING EDUCATIONAL PSYCHOLOGY — OR EDUCATING APPLIED PSYCHOLOGISTS?

Who are educational psychologists, what do they do and how worthwhile is it? Recent discussions have revealed that the "several faces of educational psychology" (Wright, 1986) have not been discriminated (Claxton *et al.*, 1986). Is professional educational psychology different from its academic relation, and, if so, is this a difference merely of setting or is there a more fundamental distinction? These questions will be considered from the perspective of a professional educational psychologist. A critique will be presented of the current interface between the two variants of educational psychology.

### REFERENCES

CLAXTON, G. *et al.* (1985). Psychology and Schooling — What's the matter?! University of London, Institute of Education, *Bedford Way Papers No. 25.*

WRIGHT, H. J. (1986). *The Several Faces of Educational Psychology.* Symposium presentation to the British Psychological Society Annual Conference, Sheffield.

NEWTON, P., and HASLAM, S., Huddersfield Polytechnic.

## GIRLS AND COMPUTERS IN SCHOOLS: A SYSTEMS FAILURE?

Computers are a scarce resource in most British schools. This often results in girls having little experience or negative experience with computers. Although several pilot projects have provided models of the creative use of computers in primary education, the available evidence suggests that computers are being used much less imaginatively in secondary schools. Computers are most frequently associated with mathematics and science subjects and are more rarely used in English and arts subjects, thus reinforcing gender stereotypes. Although the proportion of young women studying engineering has increased dramatically in the last ten years, the proportion of women studying computing has declined. The authors draw parallels between the emerging literature on gender and computing and the well-established literature on women and science. They conclude by suggesting strategies for changing the perception of computers in schools and call for radical changes in the way teachers learn to use computers.

PEARSON, L., Psychological Services, Birmingham LEA.

## APPLYING PSYCHOLOGY TO EDUCATION

This review paper provides a critical overview of what psychologists employed by local education authorities offer, illustrating this with examples of some recent work. These are then considered with reference to any proof of effectiveness and suggestions made for future developments with particular reference to the gap between professional and academic psychology and the need to narrow this.

PUMFREY, P., Department of Education, University of Manchester.

## ATTITUDES TOWARDS THE ROLE OF THE FORM TUTOR IN SECONDARY SCHOOLS

The purposes of this paper are twofold. The first is to present evidence from three recent empirical studies carried out in the Centre for Educational Guidance and Special Needs into secondary school teachers' attitudes towards the role of the Form Tutor. The second is to identify ways in which the initial and in-service training of teachers for this role can be improved.

Initially the identification of six aspects of teachers' attitudes towards the role of the Form Tutor is presented together with an indication of the amount of time spent on this work.

Changes over time in the attitudes of a representative sample of 132 teachers in a specified LEA when in training and also during their first year of teaching are described. Suggestions for improving both the initial and INSET training of the Form Tutor are analysed.

A scale designed to identify attitudes on a dimension ranging from Tutor Ascendant through Tutor Neutral to Tutor Subordinate, is also described, together with preliminary results of some between-schools differences.

RAVEN, J., Edinburgh.

## A CRITIQUE OF CURRENT PRACTICE IN TEACHER EDUCATION AND THE ROLE OF THE PSYCHOLOGIST

In this paper it will be argued that the most important contribution which psychologists can make to teacher education does not involve teaching psychology to undergraduates. It involves applying psychology to (teacher) education.

There are some fundamental problems with teacher education. These will be illustrated from published enquiries into education and from the author's interviews with college staff, practising teachers and students. These problems are of long-standing: they have been noted by committees of enquiry into teacher training over the past half century and have not been solved by their recommendations.

The problems include: failure to foster in students the competencies required to manage classrooms and maintain discipline; failure to foster the competencies required to lead pupils to understand the processes required to communicate effectively or those which lie behind mechanical arithmetic; failure to foster the competencies required to implement individualised, competency-oriented programmes of education; and failure to foster the competencies required to undertake curriculum development and innovation.

The repeatedly proffered solution is for student teachers to spend more time in schools. To be effective, this solution requires the teachers with whom students are placed to possess the competencies which are needed to foster the development of colleagues and articulate for them their own classroom management strategies. More basically still, it depends on the teachers with whom students are placed actually leading pupils to understand mathematical processes or fostering different talents in different pupils.

What have we, as psychologists, to contribute in this situation? We have something to say about the strategies which tutors can use to create developmental environments for the teachers with whom students are placed. We have something to say about how headteachers can create developmental environments for the staffs of their schools (and such staff include student teachers). We have something to say about how headteachers can create climates which stress innovation and high standards in schools. We have something to say about the kinds of research which tutors should be pressing their colleges to undertake so that they can have a better formal understanding of some of the relationships which students need to understand. We have something to say about how tutors and teachers can foster effective action strategies in others. We have something to say about how government departments can foster more developmental environments in colleges.

Most of the research needed is psychological. If more developmental environments are to be created for student teachers in schools, it will be necessary to clarify the nature of managerial (headteacher) competence — and especially the strategies which are to be used to identify and develop the talents of subordinates. It will also be necessary to identify ways in which teachers can facilitate the development of teaching competence in student teachers. It will be necessary to identify the competencies which teachers require to do their jobs effectively and how the development of these is to be facilitated. To define the qualities which student teachers are to be helped to develop, it will be necessary to clarify the competencies which are to be fostered in pupils and how those competencies are to be fostered. If teachers are to assist student teachers it will be necessary for these teachers to be able to get credit in staff-appraisal exercises for having spent time making explicit their educational theories and facilitating the growth of student teachers as well as pupils — yet, to date, we know little about how such contributions are to be assessed. And if teachers are to be able to do their jobs effectively, we will need to provide them with a range of diagnostic and prescriptive tools which will enable them to implement individualised, competency-oriented, educational programmes and take appropriate, individualised, remedial action when necessary.

It would seem from this discussion that the problems which plague the educational system will not be solved by offering teachers longer or re-vamped training courses as these are currently defined. They will not be solved by pitching student teachers into schools. They will only be solved as psychologists develop a better understanding of the educational process — for pupils, student-teachers and teachers alike — and provide the tools which are required to implement that process and testify to the outcomes. The most important role for psychologists in colleges of education is, therefore, to develop the concepts, understandings and assessment procedures which are a pre-requisite to any effective system of education.

SMITH, L., Department of Educational Research, University of Lancaster.

## DEVELOPMENTAL THEORY AND TEACHERS' PRACTICE

It is accepted in all major educational documents, from the Hadow Reports to *Better Schools*, that the promotion of children's development is one main aim of schooling. This paper examines the adequacy of developmental theory in relation to teachers' practice. The paper is in two parts:

*What is developmental theory?* Five criteria are identified: cognitive change over time; internal change with external causes; adult knowledge as indicator of developmental level; self-construction of cognitive change; incomplete construction of knowledge; complete construction of knowledge. These criteria are illustrated and discussed in relation to psychometric interpretations of development.

*Objections.* There are at least two objections to the use of developmental theory in the classroom:

*Does practice precede theory?* The objection is that good practice is evident in advance of developmental theorising in psychology. The reply is that teachers who accept the importance of developmental issues in schooling will presuppose some theory in their practice. An important empirical issue concerns the nature and the number of developmental principles and techniques which underpin the practice of good teachers.

*Does practice have priority over theory?* The objection is that theory is judged in relation to practice rather than the other way round. In reply, it is argued that even skilful practitioners benefit from the contributions made by reflective practitioners. It is unclear whether common sense embodies discrepant or several contrary or one consistent notion of development.

The conclusion (Smith, 1987) is that it is an empirical and open question whether developmental theory in psychology has a valid and useful contribution to make in the classroom.

### REFERENCE

SMITH, L. (1987), Developmental theory in the classroom. *Instructional Science* (in press).

TOMLINSON, P., School of Education, University of Leeds.

## PSYCHOLOGY IN TEACHER EDUCATION: ONCE THERE WAS A WAY . . . THROWING IT AWAY . . . DOING IT THEIR WAY

Today the role of educational psychology in teacher education is being questioned and revised as never before. This paper considers the forms in which psychology may make a contribution in the education of teachers against some of the lessons to be gleaned from a brief review of the recent past.

Psychology seems to have progressed from a paradoxically dominant role in the technical rationality of postwar teacher education to a Catch-22 role threatened by rejection and marginality in the current scene. The paradoxes include the early dominance of a discipline with very little content of relevance to the teaching process, and yet the almost total failure of educational psychology in the 1960s and 1970s to adopt and adapt the insights of the dominant paradigm in the "pure" field, whose relevance to educational and developmental issues is only now being recognised in the current decade. The reaction of calling for a return to the classroom appears to have both positive and negative potential; it is argued that we do not need total innovation for the development of a classroom psychology.

There appear to be two main ways in which educational psychology may make its contribution, though both depend for their effectiveness on our adopting what I have elsewhere termed an "open-critical" perspective and on being willing to use psychology reflexively. The first is to seek action principles and strategies of direct relevance to aspects of schooling, and to demonstrate their relevance through practical implementations and training based on them. The second is to mesh psychological perspectives into the range of sources teachers will hopefully use in their development as reflective practitioners. The thesis is that we need to be both practice-oriented and philosophically reflective in our use of psychology. Without the former we shall become irrelevant, without the latter our relevance will be merely that of short-term gimmickry and fragmented technique.

TURNBULL, H., Psychology Department, Craigie College of Education, Ayr.

## PSYCHOLOGY IN TEACHER TRAINING

The main aspects of psychology covered in teacher training are educational and developmental. Various contemporary questions on courses in psychology within teacher training centre on What? How? and When? The "What" question is about which aspects of educational and developmental psychology require to be taught, e.g., what should be the balance of cognitive development as compared with emotional development? Is "intelligence" still an important concept? Is motivation more important than social behaviour? "When" questions relate to the part of the course when psychology should be concentrated, or should it be spread evenly all through. The most difficult question is *How?* Should psychology be part of an integrated or professional studies course or should it be considered a separate subject? Should it be directed towards specific school experience placements? Should it include or be included *within* courses on Special Needs? Is Assessment best taught as a general topic or as Assessment in Science or Assessment in the primary school? Is Discipline best considered as an overall theme or as part of a psychology course?

Finally, is there a different role for psychology within different types of course? Should it be differently considered for in-service as opposed to pre-service? Is it the most important part of teacher education?

WARE, J., and EVANS, P., Institute of Education, University of London.

## SOME EFFECTS OF ROOM MANAGEMENT WITH CHILDREN WITH VERY PROFOUND LEARNING DIFFICULTIES

Although almost all published research into the effects of Room Management has been conducted with profoundly retarded adults in health or social services settings, this technique has become increasingly popular in classrooms for children with profound and multiple learning difficulties. Early adult studies (e.g., Porterfield and Blunden, 1977) suggested that Room Management was very successful at increasing clients' engagement levels. However, more recent studies have suggested that the engagement of some more handicapped clients is not increased by Room Management procedures (Mansell, Felce, De Kock and Jenkins, 1982).

The pupils in two classrooms for children with profound and multiple learning difficulties were observed under Room Management and non-Room Management conditions using both a modified version of the PLA-check and a more detailed observation schedule. To some extent the findings from adult studies were replicated, with the most profoundly handicapped children being engaged for only a small proportion of the time even when Room Management was in force. However, the use of Room Management did make a difference to the most handicapped children in ways that do not seem to have been examined by previous investigators. These differences, and their implications for children with profound and multiple learning difficulties, are discussed.

WATSON. J., Department of Special Educational Needs, Moray House College of Education, Edinburgh.

## PROFESSIONAL AND PERSONAL DEVELOPMENT IN TEACHERS OF CHILDREN WITH SPECIAL EDUCATIONAL NEEDS

The present discussion cites data from pre- and post-course interviews of one year's cohort of teachers undertaking a full-time in-service diploma course at a Scottish college of education. All were experienced teachers working in a variety of settings with pupils with special educational needs. The course is designed for teachers of "recorded" pupils.

Course contents are not described in detail. What is of interest are the expectations and hopes of course members before the course, and their retrospective evaluations of what the course had meant to them in professional and in personal terms.

Data will be presented in the following manner:

(1) characteristics of course members;
(2) their expectations and hopes concerning the course;
(3) their evaluations of personal and professional change as a consequence of the course;
(4) general conclusions.

WINTERINGHAM, D., Psychological Services, Coventry LEA.

## PROBLEM BEHAVIOUR IN PRIMARY SCHOOLS — A WHOLE SCHOOL APPROACH

Since 1982 the Coventry SNAP programme has been involved in meeting special needs in primary schools. In recent years there has been an increase in requests to support services for help with children experiencing behavioural and adjustment difficulties. As part of the SNAP initiative a group of teachers and psychologists have therefore worked on creating a package of materials aimed at producing a whole school response to behavioural difficulties in primary schools. This paper describes the development of the package and the problems it posed both in its creation and implementation.

# BOOK REVIEWS

CALAM, R., and FRANCHI, C. (1987). *Child Abuse and its Consequences: Observational Approaches.* Cambridge: Cambridge University Press, pp. 226, £7·95 pbk., ISBN 0-521-31614-6.

The detailed observation of parent-child interactions, in families where child abuse has occurred (or is suspected), is too often under-utilised as a means for understanding some of the dynamics of family violence. Calam and Franchi present their data and analyses from their own observational study of eleven (3- to 4- year-old) abused children who attended, with their nine mothers, a NSPCC Family Centre. Interview data, together with information from social work files, are also included, to provide background case study descriptions of each child and his or her mother.

Overall, the authors produce a vivid description of several aspects of abused children and of abusive parenting. There are, however, undeniable shortcomings in attempting to present an account of the consequences of child abuse without overtly including fathers. (It is such a weakness that, perhaps, the book's title should have been stated more explicitly.) Even though it is acknowledged (p. 37) that an imbalanced report will ensue, because of this omission it is surprising not to find in Chapter 12, when the authors discuss future research into child abuse, a research model which would include fathers (and any other significant (abusing?) adult). Practitioners working with child abusing families are unlikely to be wholly convinced by research until the horrors of child abuse are considered primarily as a whole family problem.

The 14 chapters are of mixed quality. When the authors are dealing specifically with their own observational and interview data they are always of relevance. At other points, especially in the latter part of the book, Calam and Franchi merely provide a very general discussion. There is an attempt at providing follow-up material in Chapter 11, but this is of questionable value (because of research constraints beyond the authors' responsibility). In Chapter 13 the authors exemplify how the provision of day care for abusing mothers and their children is insufficient per se; but this is, in the Family Centre studied, clearly only because of how little time the mothers actually spent interacting with their children. Presumably that is why there is a limit to their observational data on parent-child interactions.

Child-care professionals can gain, however, from reading this text. The observational data contained in Chapter 8 ("Approach, reaction, avoidance and hostility") illustrate specific emotional consequences of abuse to young children. Information obtained from interviewing the children's mothers, on issues of affection, discipline and control (Chapter 10) frequently expose the contradictions of their purported "good-enough" parenting. Moreover, Calam and Franchi succeed in providing detailed accounts of abused children's interactions with their peers, nursery officers, and, on a limited basis, with their mothers (as these are mostly restricted to lunchtimes in the Family Centre: see Chapter 6). There is, nonetheless, a worrying lack of observational data about how these children generally interacted with their mothers. However, interview responses to questions about bed-time routines and parents' understanding of their children's sexuality (Chapter 9) provide information which should lead others to further these insights into what is a poorly understood area of parent-child relationships.

Even though Calam and Franchi's overall research design may have limitations, ultimately because of a lack of opportunity for more extensive observation of how mothers *and* fathers interact with their children, they have succeeded in showing the value of how an observational study can give some important additional insights into distinctive features of abused children's behaviour. Infant and pre-school teachers, nursery officers, social workers and child psychologists/psychiatrists can surely be included in those already listed who will find much of this book interesting and potentially useful.

NEIL HALL

Edwards, D., and Mercer, N. (1987). *Common Knowledge.* London: Methuen, pp. 193, £9·95 pbk., ISBN 0-416-39900-2.

*Common Knowledge* is a nicely titled study of how meanings are imparted, negotiated and shared within classrooms. Based on the analysis of selected videotaped lessons with groups of 8-10 year-olds, supplemented by material from interviews with their teachers, this book should take its place among a considerable number of Open University set texts about language and literature teaching that have informed and edified a much wider audience than just OU students.

Few people answering the question "Where's the railway station?" would expect the response "Well done!" or "Right!" People who know do not normally ask questions. Only teachers — or parents perhaps. Douglas Barnes was among the first to draw attention to the complicated nature of classroom transactions: it wasn't a simple matter of pupils learning what Mr X (thought he) was teaching but rather of learning what it was to be a learner, of a particular subject, with a particular teacher, in a particular place, on, say, a Friday afternoon.

This notion, and that of context of utterance, are not new — Tolstoy evoked them a hundred years previously in *Yasnaya Polyana* — but their implications are still being worked out. Following Hymes, Cazden and others, Bernstein finally pronounced: "The particular form of a social relation acts selectively on what is said, when it is said, how it is said". There are simpler ways of putting it, and one of the virtues of this book is its clarity but the authors go further. They extend the notion of context to cover not merely the social dimension but the whole panoply of shared meanings and how these are arrived at in the classroom negotiations between teacher and taught.

As I said earlier, it is all very complicated — "problematical" the authors would say. Indeed the word "problematics" occurs often enough to jar on a reader who feels that the language would be richer without it. But this is exceptional. Their exposition of knowledge and its tricky relation with both the teacher's conception of his role and the culture, the meanings, which he is expected to impart is modest, strenuous, generally elegantly expressed, and sustained.

"Sustained" — there indeed lies a slight rub. As the authors make clear, the evidence for the study is deliberately limited but the extended analysis poses some problems for the reader. The first is general: even when subtly explicated, transcripts are not easy bed-time reading. Secondly, where a limited range of transcript data are analysed from several points of view the reader may be bothered by the feeling that he has been here before, or that he is sucking an orange that by now seems somewhat dessicated.

These are not criticisms. If your model of research is one of careful exploration and speculation as opposed to the derivation of prosaic conclusions from mountains of data, those risks have to be run and the authors defend their approach convincingly. I am particularly grateful to them for avoiding over-simplification, especially one which has dogged some writing on the subject: that it all depends on whether we're looking at "progressive" or "traditional" teaching. A little depends, perhaps, but what they bring out so clearly are the elements that are endemic to the classroom: the real differences are between pedagogic and non-pedagogic settings, not between teaching styles.

I conclude with another note of gratitude — for their second chapter, "Approaches to Classroom Knowledge and Talk". At a time when Bullock, let alone Plowden, seems to have been buried, when reform is in the air, when floods of slogans masquerade as educational ideas, it was good to revisit trustworthy old friends. For the students who here will be meeting them for the first time, this review of approaches, which often underlie and inform the rest of the text, could hardly be bettered.

Paddy Creber

LANGFORD, P. (1987). *Concept Development in the Secondary School.* Beckenham: Croom Helm, pp. 172, £19.95 hbk., ISBN 0-7099-4163-3.

Peter Langford has written a book which is to his credit. He deals with the relatively neglected areas of intellectual development relevant to secondary schooling. The book covers a range of school subjects: English, Humanities, Social Studies, Science and Mathematics. Such breadth in one volume from one author is rare. Full and recent source references are given in each of these areas — even advanced readers will see the benefits here. This book could profitably be on the shelves of any institutional library. The author is an inspiration to us all; yet the book has two shortcomings.

One shortcoming is that practitioners may wonder what they are being given. Consider: "the aim of this book is to introduce the secondary teacher to recent work in the area of cognitive development that is relevant to the interests of the classroom practitioner" (p. 1). Construed as a review of recent work, this is accurate. But it is also misleading in two respects. Firstly, whilst developmental research is — and is extensively — used, it is not clear what research is developmental. The title refers to concept development; the statement of intent refers to cognitive development; the last textual page makes it clear that non-developmental research also has its place: "theories of Gestalt psychology and information-processing theory can be applied to problem solving in secondary school mathematics" (p. 133). Certainly, the chapter on English draws freely from such perspectives.

A point of substance arises here, namely the criteria to be used in the demarcation of research which is *developmental.* This links to a related point that teachers have an interest in intellectual *development.* What criteria should be used to identify when an increase in learning amounts to progress in learning? Conceptual clarity is as necessary as it is welcome. However, clarity may not be enough, since practitioners also value utility. We are told, in the section on History, that "the bare outline of the development of historical thinking given here will be difficult for many teachers to relate concretely to classroom history lessons" (p. 65).

This conclusion can be generalised to each chapter. It is clear that an introduction has been given to the practitioner who has intellectual interests but not to one who wishes to know how recent research can be applied. Models of teaching, of classroom learning, of instructional procedures, of diagnosis and assessment are simply passed over in silence. This omission is crucial. The common sense which practitioners can be relied upon to use is not normally sufficient to allow them to operationalise for classroom use developmental principles generated by researchers.

The second shortcoming is that the introductions are sometimes faulty. Here is a test: name the author of this developmental theory? Its author had a cause (p. 2); structural development is maturational (p. 3); age and stage are criterially related (p. 5); formal operations consist in thinking about thinking (p. 11) such as reflecting on processes of logical deduction (p. 132); development is marked by sudden transitions (p. 14). Langford uses two French texts by this author (one of which was translated in 1974) but passes over another dozen recent volumes in silence. Relevant tasks are randomly reviewed. In the chapter on Science, seven formal operational tasks are mentioned as if they were severally independent. Another pops up in the Mathematics chapter after its debut in the History chapter. You can think what you want of these tasks but the unitary nature of the concomitant theory is lost in this discussion. This is an introduction not so much to the giant of developmental psychology as to the phantom of the secondary literature.

Here is another test: with respect to history learning, what are the similarities, what are the differences, between the positions of Hallm, Dickinson & Lee, Shemilt and Booth? Information is provided in the text on these several positions — but you would have to search your imagination to make sense of such listings.

Entries to the comprehensive bibliography are arranged by chapter at the end of the book. Evidently development is modular rather than unitary.

LESLIE SMITH

McNaughton, S. (1987). *Being Skilled. The Socializations of Learning to Read.* London: Methuen, pp. 276, £32·50 hbk., ISBN 0-416-016227.

This book presents a model of learning to read which brings together, in a single perspective, accounts of children who are able to read before they come to school and those who learn to read in school. It focuses on high progress readers, that is those who learn to read early, easily and well and maintain this ability. (Precocious readers almost always stay within this group.) McNaughton examines what such children actually do as they learn to read and the environments in which they operated. In his introduction he sets the scene for his analysis. He conceptualises reading as a complex skill and readers as problem solvers whose effectiveness is dependent on their intentions to learn and the control they have over the learning. The environment and socialisation within which learning takes place is seen to be composed of variables which determine the intentions of learners and the sorts of activities they engage in.

The author seeks explanations of three features of learning to read which he states must be addressed in articulating a theory of reading which takes into account "naturally occurring phenomena". Drawing widely on existing research he details these features. The first he describes as "inventive conformity". This relates to the failure of research to demonstrate the superiority of any one method of instruction over another. Children can and do learn to read whatever the programme of instruction. The second feature "going beyond the teaching received" relates to research showing that children invent elements of reading skill that are not included in particular programmes, which they would have been taught deliberately by an alternative method. For example, they learn phonics in a natural language programme. The third, "never schooled yet learned", is the phenomenon of children who read precociously. Sometimes such children are taught, more often they initiate the process often to the embarrassment of parents who nevertheless respond. Such early reading does not necessarily demand high social class or high tested intelligence.

What follows is a densely argued confrontation with these issues. It draws extensively on existing research including the author's own. Ideas from both child-oriented perspectives and the various types of learning theory are included and amalgamated. The model presented is that of construction: readers solving problems presented by the learning task aided or hindered by the dynamic environments which they construct jointly with their tutors.

It is argued that reading is a strategic skill, that performance strategies derive from knowledge and that this knowledge is the basis of self-regulatory strategies. Strategies involve conformity to rules and making choices but will vary according to knowledge which is acquired through socialisation. Action and its regulation are said to be achieved through intention, feedback and intervention. These characteristics are seen as crucial and are returned to and examined in detail throughout. Intention is central to the argument as it is seen as setting conditions for effective learning and allowing for "performance directed activities". Although "performance directed" regulation, that is, strategies under the control of the child, require information from skilled owners it is argued that such assistance may become interference. Unfortunately interference undermining self-regulation seems to be a common experience of low-achieving children.

Many other issues are examined from different perspectives. Skill in development, particularly learning from dissonance, modularisation (increasing automaticity and fluency), and the restriction of performance directed learning by external dominance of tutors are discussed. The author describes the variable socialisations which children encounter as they become literate. Unsurprisingly, he argues that when school is a very different social environment from home, children are more likely to be low achievers. Children who have not learned to read before they come to school are just as likely to be high achievers as those who have, if their social experience matches that in the school.

McNaughton concludes with "tinkerings" which might help, for example by pre-school teachers being seen to value reading by reading at their own level in front of their classes. Smaller classes are also recommended. This hardly seems to be a "tinkering" but an alteration of the political will unlikely to come about here, although the author is from New Zealand which may be more forward looking. Or it may have been realised there that the body of research indicating that smaller classes are not productive relates to a rather narrower range of skills than it is actually demanded that teachers shall teach. However, his major conclusion is that real change will only come about if political changes occur to transfer control of educational policy to parents and if they are enabled to see reading as powerful in social, economic and political terms. This is also in the context of New Zealand and its indigenous population but may have some significance elsewhere.

Amongst interesting questions which this book raises in the mind of the reader is what might be done to encourage all children to go beyond the immediate demands of a particular learning programme or task. The contention that low achievers have what strategies they do use consistently interfered with raises questions not only about how to identify the point where feedback becomes interference but also, and more generally, how the sub-cultures of schools might adjust to the different socialisation experiences of their children. If, as McNaughton argues, schools are not very effective situations for individual learning, especially for children whose socialisation does not predispose them to develop their own strategies, what effect is the current press for testing likely to have? Will it lead to less focus on the child as problem solver and more on mechanistic skills, putting more difficulties in the way of less able readers?

In general this book opens up a fresh dimension in theories about learning to read. The arguments given are detailed, well documented and persuasive. This book is likely to be very useful to its target audience of advanced students, researchers and theorists in education and developmental psychologists. If these in turn have anything to say to teachers and teachers of education they will have to address the author's wry comment: "The harsh fact is that one to one interactions are counter-productive because of how classrooms need to be managed". This is immediately followed by his own argument that a child is most likely to learn in a one-to-one setting. McNaughton's readers will also have to address the social and political implications of his theory.

DEIRDRE PETTITT

PARSONS, M. J. (1987). *How We Understand Art.* Cambridge: Cambridge University Press, pp. 159, £19·50 hbk., ISBN 0-521-32949-3.

Michael Parsons's *How We Understand Art* has been long and eagerly awaited by those of us who have followed his investigative work, available hitherto only as individual, usually collaborative, papers, published in learned journals at intervals over the past 10 years or so. Serious research in the field of aesthetic development is notably scarce: how we understand art is every bit as "contested" a concept as art itself, and studying the operation of sensibility poses acute methodological difficulties to researchers of every persuasion. Added to which, unlike the areas of scientific and moral reasoning, the aesthetic area has always lacked the necessary sense of political and pragmatic immediacy. And yet, of course, for art educators in the present climate the need for an adequate theory of aesthetic development was never perhaps so keenly felt, both as reinforcing the case from the continued presence of the arts subjects in the general curriculum and as affording a conceptual framework to inform practice and guide evaluation.

In such circumstances it is therefore all the more regrettable to have to admit to being disappointed with this book. Not that the study is inherently understanding — there is much to ponder and debate — but that it seems wrong-headed. One senses that Parsons's pre-occupation with the way people think about pictures is no longer — if it ever was — the burning issue in *arts* education. Scarcely, even in art education — he himself confesses to having had his own doubts on the subject. Nowadays we need to know about *the arts as a way of knowing* rather than *ways of knowing about the arts.* Parsons's investigation, undertaken and presented in a disarmingly relaxed and informal way, has rather been marginalised by the progress of events — in education itself, as well as in the world at large. Under the pressure of fresh imperatives coming from the politics of social anthropology on the one hand and a post-modern sensibility on the other, arts educators are now actively seeking a radical new concept that would lend relevance and force to their work — not least a concept to make practical sense of the principle of an educated aesthetic response that goes beyond the notion of "art appreciation". Parsons's study says nothing in this connection.

Parsons proposes to do for the aesthetic area what Piaget and Kohlberg have done for our understanding of the child's conceptual and moral development. This in itself invites some scepticism in view of the enormity of such an undertaking. That scepticism converts to distrust when one realises that the research rests exclusively upon an examination of verbal statements about a group of pictorial reproductions. It just does not seem credible to propound a theory of aesthetic development (cognitive or otherwise) on the basis of such a partial and highly selective enquiry — one that furthermore never attempts to examine the distinctive and separate problems that flow from the decision to rely solely upon verbal data.

Parsons's 10-year study is a best speculative: he offers a hypothetical, stage-based model, following the example of his mentors, that is "tested" by means of an analysis of what some 300 assorted persons living in and around Salt Lake City, ranging in age and experience from infants to professors of fine art, have to say about a handful of paintings culled from the mainstream European high-art tradition, c. 1900-1950. For "art" read therefore this highly selective material; for "aesthetic experience" read similarly. The study makes claims for itself that go way beyond what it can deliver — indeed these claims tend to reveal rather damagingly the study's adherence to some very narrow and now dated notions concerning arts education, possibly still relevant to some degree in the US, but in my opinion of marginal interest in the UK. And that despite the current questionable vogue for the wholesale revival of critical studies.

What do we get? A Kohlbergian model that purports to distinguish five stages in the development of the aesthetic response. These stages are roughly aligned with chronological age and reveal basic changes in what we expect of an art object, in our ideas about what a painting is, what purpose it serves, how it works, what is has to offer. The stages are negotiated sequentially, some people never achieving the level of an ultimately "fully adequate" aesthetic response: the progress is the typically cognitive one from ego-centricity to allo-centricity, from concrete to abstract thought, from feeling-dominated to idea-dominated judgment. The subjects are asked to *describe* the painting, to say what it is *about,* to identify the *feelings* in the painting, to say whether the *colours* are good, to assess its *texture,* its level of *difficulty* and finally to say whether or not it is a *good painting* and why. Responses range from "It looks like licorice" to "It's very traditional, excessive chiaroscuro, light logic". The hypothetical model just seems too neat: having reasonably identified four key concepts basic to an analysis of the notion of art — content, expression, medium and judgment — Parsons suggests that each in turn provides the characteristic emphasis of a developmental stage. The book strains to endorse this plausible but hardly compelling thesis.

Some years ago the philosopher, the late Professor Louis Arnaud Reid, reviewing Parsons's work appreciatively, felt nonetheless that although on the right lines it was not radical enough. Reid suggested that in the work of the "cognitive school" as a whole the "continual and almost exclusive emphasis (in the discussion of children's development) on *art*" was overdone. What of the world outside the arts? he asked. Aesthetic education has a broader and more urgent agenda than Parsons attempts to address. This book, alas, is not going to be much help in our present parlous situation.

MALCOLM ROSS

---

PELLEGRINI, A. D. (Ed.) (1988). *Psychological Bases for Early Education.* New York: Wiley, pp. xii + 282, £29·50 hbk., ISBN 0-471911453.

This is an interesting collection of papers well worth a place on the shelf of any institution educating teachers of young children. The least impressive facet of the volume is its theoretical rationale. The editor proceeds on the assertion that early education should, indeed must, be based on psychological research on young children. There is no exploration of the problems of taking this stance — but no matter.

The second least interesting thing about the collection is the exhortation to consider the whole child and to avoid focusing on disintegrated dimensions of childhood such as cognition or temperament or socialisation. Several of the contributors strive mightily to live up to this exhortation. Most notably, Case offers a potted version of his general theory of child development drawing from social and cognitive factors in building a model positioning powerful, general dialectical forces to be at the heart of psychological growth across the board.

However, much the most interesting material to this reviewer ignores the editor's stance on integrating academic disciplines and dimensions of childhood. Several authors offer reviews or reports of recent research on children's intellectual growth in substantive areas in the early childhood curriculum. Included here are chapters by Sulzby on the precursors of reading and Martlew on early competence in oral and written language. Each is worth the asking price of the book. But selected for special commendation must be the chapter by Wolf and his colleagues in which they argue the case for a broader and deeper view of literacy.

In this context the authors refer to literacy as the capacity to use any symbol system to represent experience. Breadth in literacy refers to the range of notation available including, for example, numbers, letters, musical script, diagrams, maps, drawings and the like. Their notion of depth refers to the processes of selection which abstract from experience those factors thought necessary to be recorded. Depth also refers to the processes of organisation which guide the way in which symbols are used to indicate the substance of experience.

The authors report data from longitudinal studies of children age 5 through 7 years indicating their growth of competence in idiosyncratic symbol systems for layout, music and number. The research is extremely provocative and challenging. The children show remarkable invention in these domains and powerful skills in selection and organisation. The authors cannot resist comparing their powers with those typically demanded by early childhood literacy curriculums. These available curriculums are designed on the basis of task analyses which at heart contain assumptions about the nature of literacy and about young children's competencies. Wolf and his colleagues show how these assumptions make monkeys out of children. This message is not a lot of use to teachers trapped by the curriculums imposed by their masters. But Wolf's paper should be compulsory reading for all those who have got fat on the profits from writing schemes for infant and nursery schools. More profitably perhaps, we can look forward to the development of Wolf's programme into the field of curriculum design through research.

CHARLES W. DESFORGES

---

REYNOLDS, C. R., and MANN, L. (1987). *Encyclopedia of Special Education, Vols. 1, 2 and 3*. New York: John Wiley & Sons, pp. 1793, n.p., ISBN 0471-82858-0.

In reviewing a set of encyclopedia, one is faced with a number of problems — inability to know with any level of familiarity more than 5 per cent of the contents, lack of time to read more than 10 per cent, and the motivation to "read" an encyclopedia at all. However, the authors have succeeded in engaging my interest. What I did read was clearly written and understandable in my opinion to a wide range of readers from students to lecturers. I tackled the sheer volume of the books by referring to a number of items I thought I had some level of knowledge about, and also a range of items I felt I knew little about. This was I felt a fair way to tackle the project.

The authors list four main uses for the encyclopedia:

(1) As a source of basic summative information . . . and readers seeking information will find it available in quick and easy to comprehend form, whether the information they seek concerns definitions, concepts, theories or historical personages. . . .

This claim I feel is upheld: there is clear, easy to comprehend information about a whole range of topics, from those that deal with issues specific to special education to those that would also be relevant to mainstream education.

The authors also suggest the encyclopedia can be used "As a basic spot reference . . . using the entries or extensive index, the user can look up and check out particular topics for names, definitions".

I feel here the authors undervalue what they have compiled. The way the entries are cross-referenced would lead the novice to a host of alternatives and linked topic areas which would immediately broaden their perspective. In addition, each *main* item has at least two recent references which would be useful as starting points for further reading. My only regret is that the items without references are the less commonly referred to, items in general literature, and would benefit from exposure and elaboration in such an enterprise.

(3) The third point the authors make is that "the encyclopedia is not intended as a primary text for basic or core courses. . . . However it will serve as a useful supplementary text for almost any course . . . can serve as a primary text for advanced courses; as such, the student will be left with a set of books that will serve him or her long into a professional career."

The implication as I read it was for students themselves to purchase the texts. This I feel would be unrealistic on a number of counts. The encyclopedia covers at a superficial level a wide range of topics. The inference that the encyclopedia would serve the student long into their professional career is a doubtful one. The student may well have specialised beyond the content of the text, and there are many items which will need frequent updating — areas such as tests and testing, checklists and those entries involving computers. The encyclopedia, produced in USA, has of course been written from an American perspective and this limits it usefulness to non-American readers. Although there is much of use, certain items are of less critical interest.

(4) The fourth purpose the authors suggest for the encyclopedia is "As continuing information-support system. . . . Well beyond the immediate needs of knowledge and information the encyclopedia will serve as a source of continuous education for its owners. Regular, even review of its contents will ensure its readers a deepest understanding related to special education and related fields."

I have some disagreement over this claim. Many entries will be of only relative interest for non-Americans, and to American students there will also be the question of contemporary relevant of the entries and their supporting references. Much of the information will remain stable over time but those I have mentioned under point (3) (and many more) will decay with time, making regular review of these items pointless.

*Main disadvantages*

(a) There are several entries which would be of *prime* concern to American readers, or those interested in the American system (e.g., Public Law 94-1422 which takes up some 18 pages, and continual reference to DSM III — Diagnostic and Statistical Manual III mainly available and familiar to practising professionals in USA).

(b) There are omissions of several low frequency syndromes (e.g., Cordis syndrome, Treacher-Collins syndrome). I tried cross-referencing to see if there were different American labels for these syndromes, but couldn't locate any information. I would find entries on rare syndromes a particularly useful reference for such an encyclopedia.

(c) My third concern is the "life span" of such a text. Education and special education seem to be undergoing a period of rapid change and I wonder how relevant some of the entries will be in, say, three years time.

*Main advantages*

(a) A useful starting point to access information across a wide range of related topic areas that are clearly cross-referenced.

(b) A good source of topic related references that will guide readers to further books and articles. As a tutor I will find it particularly useful as a reference for essay topics.

(c) Some good sections on wide ranging areas such as Arithmetical Retardation and Trisomy 21 (Down's syndrome). These entries address topics spanning vast areas and manage to give clear and relevant information and key points. This would help the reader focus on the key issues and move forward to further reading, avoiding getting overwhelmed at the beginning of the information search.

(d) There are references to what might be considered Americanisms (e.g., ADD Attention-Deficit Disorder). However, I find these useful when added to the British perspective about the same topic areas.

*Usefulness*

(1) To students — useful for starting points, useful references and clarification.

(2) To lecturers/tutors — useful set of books to access for themselves and to recommend to students.

(3) To schools and teachers — useful for clarification, and for starting points. There may be some confusion, however, when tests are mentioned. There is no mention of which tests are open and closed to teachers, and it may raise spurious expectations.
(4) Parents — if the encyclopedia were lodged in the public library, parents might find it useful to translate what various professionals have said to them about their child.
(5) Libraries, colleges, schools — I would think that this encyclopedia would be useful even though the life-span of certain entries may be limited.

TERRI-ANNE HORNBY

---

WADE, B., and MOORE, M. (1987). *Special Children . . . Special Needs.* London: Robert Royce, pp. 212, £10·95 pbk., ISBN 0-947728-36-8.

An excusable first reaction to this book would be to criticise spending nearly two-thirds of the text on nine case studies intended to illustrate nine areas of special needs. Of what relative importance are these chosen areas? Are they sufficiently representative? Further, are the individual children chosen well, as we all know that there isn't one type of disruptive child or one type of reading difficulty. Hence my early anxiety.

The chosen areas do not include the usual traditional "categories" of special educational need such as physical handicap or visual impairment but do include "the abused child" and "the child of exceptional ability". In reality the latter two areas are not so readily noted as the more obvious physical handicaps or visual impairments. Further, although beyond the brief of most similar textbooks, and indeed of the Warnock Committee, it can only be healthy to widen the concept of special needs. That SEN children can also sometimes have exceptional ability or be victims of abuse also needs to be more widely recognised.

The sheer quality of the nine case studies even managed to overcome my deeply ingrained prejudice about "typical case examples" as all too often these are ill-chosen, sometimes inclined to naïve niceties and generalised solutions, even occasionally a miracle cure! However, none of the nine case studies is guilty of any of these misdeeds and neither was it possible to offer an honest contemptuous sniff at interpretations made or advice offered.

Whilst an illustration of a disruptive child is commonly offered this is too infrequently followed by an example of a quiet or withdrawn child, as in this case.

Difficulties in reading and writing are well covered, that is, in the space available though this is hardly surprising as this is the area of speciality of both authors, or so it strongly appears. What is lacking is attention to difficulties generalised in other areas of the curriculum.

The section "organising for special needs" occupies about 30 pages and does not balance the scope of the case studies, though again what is written is well presented and chosen with care. A final selection on resources will be useful to many practititioners and parents.

When presenting special educational needs to B.Ed. and PGCE students I have always experienced great difficulty in choosing a central text, and this book by Wade and Moore is the closest yet to this idea.

MICHAEL I. SMITH

---

WESTWOOD, P. (1987). *Commonsense Methods for Children with Special Needs.* London: Croom Helm, pp. 146, £7·95 pbk., ISBN 0-7099-5247-3.

The number of new publications aiming to be an introductory text on special educational needs has reached nothing short of epidemic proportions in the last year or so, many deserving the fate of second rate popular music — soon to be forgotten.

A common recipe is to select a handful of aspects of special needs (notice how frequently the "educational" bit is now being dropped) and string a book together. What god-like omnipotence to forward the proposition that "this" will represent special needs!

Peter Westwood's book is 146 pages long and, for example, the chosen areas, eight altogether, include "helping children develop self-management skills" in 10 pages, "the management of deviant behaviour" in 13 pages and so on. Perhaps the worst example is the final chapter, 11 pages on "school-based" and regional support systems.

As with other texts, as mentioned earlier, this attempts too much within too little space. The result is superficial coverage of a few areas of special needs. If a student were to read this book it would provide a brief beginning awareness and leave too much unanswered. Of its type it is reasonably well written and aims to demystify but I'm not sure where the readership will be. Certainly not from amongst experienced teachers or educational psychologists.

The print and presentation of the book is very good, unlike a few recent Croom Helm books (have they now learnt?) and the price is fairly reasonable at £7·95.

What is worth recounting is the quotation by Johnson and Johnson that the author selects: "Placing handicapped students in the regular classroom is the beginning of an opportunity. But like all opportunities it carries with it the risk of making things worse as well as making things better. . . ."

MICHAEL SMITH

INTERNATIONAL CONFERENCE ON SEVERE AND MULTIPLE DISABILITIES
Southern Cross Hotel, Melbourne, Victoria, Australia

## CONFERENCE STREAMS

| | |
|---|---|
| STREAM ONE: | The nature of conditions with severe multiple disabilities such as cerebral palsy, multiple sclerosis and severe head injuries and on programmes and services for children and adults with these disabilities. |
| STREAM TWO: | The advances in high technology which could enhance the lives of people with severe disabilities from simple switches to battery operated educational devices, microcomputers and communication aids. |
| STREAM THREE: | Clients'/consumers' rights to have a say in their lives, and in the services provided for them; the need for some people to have advocates to speak for them. |

## CONFERENCE PROGRAMME

International keynote speakers will head up each of the streams and the conference will also feature significant Australians working in the field. The keynote addresses will include:

- **"The Process of Bringing Conductive Education out of Hungary" — Dr. Andrew Sutton.**
- **"High Technology — A Major New Tool and A Major New Danger" — Dr. Gregg Vanderheiden.**
- **"The Integration Imperative: A Vision For The Future" — Dr. Douglas Biklen.**

**The programme will include presentation of papers, workshops, trade exhibition, social events etc.**

## CALL FOR PAPERS

Proposals of papers, over a broad range of topics of interest particularly related to the three streams of the conference, are being solicited from consumers, parents, advocates, clinicians, researchers, educators and students; members of other health and education professions and interested people.

Proposals for workshops are also being sought.

ABSTRACTS: should not exceed 300 words; should state the topic or purpose of paper, nature of the presentation (lecture, panel discussion, practical workshops, etc.), desired length of presentation (10, 20, 50 or 80 minutes for conference papers, half or full day for pre-conference workshops). For detailed information on Abstract Format contact: **Convener, Program Planning Committee, I.C.S.M.D., Spastic Society of Victoria, P.O. Box 381, St. Kilda, Victoria, 3182, Australia. Phone: (03) 537 2611, Fax: (03) 525 3274.**

*Deadline for Abstracts: 15th June, 1988.*

Notice of wish to attend: Public Relations Department, Spastic Society of Victoria, at address given above.

# PUBLICATIONS RECEIVED

Books for review have been selected from the following
("n.p." means "no price mentioned").

Barber, P. (1988). *Applied Cognitive Psychology.* London: Methuen, pp. 223, £10·95 pbk., ISBN 0-416-08762-0.

Broudy, H. S. (1988). *The Uses of Schooling.* London: Routledge, pp. 131, £22·50 hbk., ISBN 0-415-00176-5.

Brown, R. I. (1988). *Quality of Life for Handicapped People.* Beckenham: Croom Helm, pp. 308, £25·00 hbk., ISBN 0-7099-3992-2.

Bruner, J., and Haste, H. (1987). *Making Sense — The Child's Construction of the World.* London: Methuen, pp. 204, £7·95 pbk., ISBN 0-416-92490-5.

Buckley, S., and Sacks, B. (1987). *The Adolescent with Down's Syndrome — Life for the Teenager and for the Family.* Portsmouth: Portsmouth Down's Syndrome Trust, pp. 165, n.p., pbk., ISBN 0-900234-19-9.

Butler-Por, N. (1987). *Underachievers in School.* Chichester: Wiley, pp. 187, £19·95 hbk., ISBN 0-471-91109-7.

Carter, F., and Cheesman, P. (1987). *Anxiety in Childhood and Adolescence.* Beckenham: Croom Helm, pp. 129, £19·95 hbk., ISBN 0-7099-4806-9; £8·95 pbk., ISBN 0-7099-4846-8.

Coles, G. (1987). *The Learning Mystique.* New York: Pantheon, pp. 330, $22·95 hbk., ISBN 0-394-54898-1.

Coltheart, M., Patterson, K., and Marshall, J. C. (1987). *Deep Dyslexia.* London: Routledge and Kegan Paul, pp. 490, £12·95 pbk., ISBN 0-7102-1235-6.

Coupe, J., and Goldbart, J. (1987). *Communication Before Speech.* Beckenham: Croom Helm, pp. 129, £16·95 hbk., ISBN 0-7099-4841-7; £7·95 pbk., ISBN 0-7099-4848-4.

Cox, K., and Desforges, M. (1987). *Divorce and the School.* London: Methuen, pp. 153, £7·95 pbk., ISBN 0-416-04832-3.

Desforges, C., and Cockburn, A. (1987). *Understanding the Mathematics Teacher.* Basingstoke: Falmer Press, pp. 167, n.p., pbk., ISBN 1-85000-213-4.

Egan, K., and Nadaner, D. (Eds.) (1988). *Imagination and Education.* Milton Keynes: Open University, pp. 288, £8·95 pbk., ISBN 0-335-15836-6.

Eisenberg, N., and Sayer, J. (1988). *Empathy and Its Development.* Cambridge: Cambridge University Press, pp. 406, £32·50 hbk., ISBN 0-521-32609-5.

Fraser, B. J. (1987). *International Journal of Educational Research Vol. 11, No. 2 Syntheses of Educational Productivity Research.* Oxford: Pergamon Press, pp. 107, n.p.

Gillham, B. (1987). *A Basic Attainments Programme for Young Mentally Handicapped Children.* Beckenham: Croom Helm, pp. 84, £8·95 pbk., ISBN 0-7099-3862-4.

Greene, J. (1987). *Memory, Thinking and Language.* London: Methuen, pp. 187, £5·95 pbk., ISBN 0-416-33800-3.

Gruneberge, M. M., Morris, P. E., and Sykes, R. N. (1988). *Practical Aspects of Memory: Current Research and Issues, Vol. 1 Memory in Everyday Life.* Chichester: Wiley, pp. 568, £35·50 hbk., ISBN 0-471-91234-4.

Gruneberge, M. M., Morris, P. E., and Sykes, R. N. (1988). *Practical Aspects of Memory: Current Research and Issues, Vol. 2 Clinical and Educational Implications.* Chichester: Wiley, pp. 571, £35·50 hbk., ISBN 0-471-91867-9.

Hallahan, D. P., and Kauffman, J. M. (1987). *Exceptional Children.* Hemel Hempstead: Prentice-Hall, pp. 530, £17·95 pbk., ISBN 0-13-293762-X.

Horowitz, R., and Samuels, S. J. (1987). *Comprehending Oral and Written Language,* San Diego: Academic Press, pp. 411, $42·00, ISBN 0-12-356-110-8.

Howlin, P., and Rutter, M. (1987). *Treatment of Autistic Children.* Chichester: Wiley, pp. 299, £29·50 hbk., ISBN 0-471-10262-8.

Hurrelmann, K., Kaufmann, F., and Losel, F. (1987). *Social Intervention: Potential and Constraints.* Berlin: de Gruyter, pp. 399, 120DM hbk., ISBN 3-11-011256-6.

International Test Commission (1986/87). *Bulletin of the International Test Commission Nos. 23 and 24.* France: International Test Commission, pp. 105, n.p., ISBN 0379-2439.

Jones, L., Kyle, J., and Wood, P. L. (1987). *Words Apart.* London: Tavistock, pp. 267, £35·00 hbk., ISBN 0-422-60970-6.

Kurtines, W. M., and Gewirtz, J. L. (1987). *Moral Development Through Social Interaction.* Chichester: Wiley, pp. 362, £30·50 hbk., ISBN 0-471-62567-1.

Kutnick, P. (1988). *Relationships in the Primary School Classroom.* London: Paul Chapman, pp. 251, £8·95 pbk., ISBN 1-85396-003-9.

KYLE, J. G., and WOLL, B. (1988). *Sign Language*. Cambridge: Cambridge University Press, pp. 318, £10·95 pbk., ISBN 0-521-35717-9.

LANGFORD, P. (1987). *Concept Development in the Secondary School*. Beckenham: Croom Helm, pp. 172, £19·95 hbk., ISBN 0-7099-4163-3.

LAWRENCE, D. (1988). *Enhancing Self-Esteem in the Classroom*. London: Paul Chapman, pp. 87, £7·95 pbk., ISBN 1-85396-000-4.

LIEBERMAN, P., and BLUMSTEIN, S. (1988). *Speech Physiology, Speech Perception and Acoustic Phonetics*. Cambridge: Cambridge University Press, pp. 249, £25 hbk., ISBN 0-521-30866-6; £38·95 pbk., ISBN 0-521-31357-0.

LONDON INSTITUTE OF EDUCATION (1987). *Report of Activities 1981-1986*, pp. 186, n.p.

MACLEAN, D., and GOULD, S. (1987). *Helping Process*. Beckenham: Croom Helm, pp. 171, £22·50 hbk., ISBN 0-7099-4682-1.

MCARTHUR, T. (1988). *Unitive Thinking*. Wellingborough: Aquarian Press, pp. 144, £5·99 pbk., ISBN 0-85030-621-3.

MCNAUGHTON, S. (1987). *Being Skilled — The Socialisations of Learning to Read*. London: Methuen, pp. 280, n.p. hbk., ISBN 0-416-01622-7.

MANGAL, S. K. (1987). *Statistics in Psychology and Education*. Delhi: Tata McGraw-Hill, pp. 244, n.p., ISBN 0-07-451629-9.

MARSH, H. W. (1987). *International Journal of Educational Research Vol. 11, No. 3 Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research*. Oxford: Pergamon Press, pp. 135, n.p.

MORTIMORE, P., *et al.* (1988). *School Matters*. Wells: Open Books, pp. 310, £7·95 pbk., ISBN 0-7291-0194-0.

OUVRY, C. (1987). *Educating Children with Profound Handicaps*. Kidderminster: BIMH, pp. 212, £10·95 pbk., ISBN 0-906054-58-3.

QUINTON, D., and RUTTER, M. (1988). *Parenting Breakdown*. Aldershot: Avebury, pp. 270, £25·00 hbk., ISBN 0-566-05582-1.

RAPOPORT, R. N. (1988). *New Intervention for Children and Youth*. Cambridge: Cambridge University Press, pp. 275, £27·50 hbk., ISBN 0-521-34122-1.

ROSENBERG, S. (1987). *Advances in Applied Psycholinguistics*. Cambridge: Cambridge University Press, pp. 326, £45·00 hbk., ISBN 0-521-30027-4; £15·00 pbk., ISBN 0-521-31732-0.

SALMON, J. F. (1987). *Training for the Management of Special Educational Needs*. Bristol: NCDSMT, pp. 73, n.p., ISBN 1-870869-001.

SIEGEL, H. (1988). *Educating Reason*. London: Routledge, pp. 191, £22·50 hbk., ISBN 0-415-00175-7.

SMITH, P., and COWIE, H. (1988). *Understanding Children's Development*. Oxford: Basil Blackwell, pp. 417, £9·95, ISBN 0-631-15723-9.

TAEUBER, R. C. (1987). *International Journal of Educational Research Vol. 11, No. 4, Education Data System Redesign*. Oxford: Pergamon Press, pp. 125.

TOPPING, K. (1987). *The Peer Tutoring Handbook*. Beckenham: Croom Helm, pp. 123, £16·95 hbk., ISBN 0-7099-4348-2; £6·95 pbk., ISBN 0-7099-4353-9.

WIENOLD, G., *et al.* (1985). *Lehrerverhalten Und Lernmaterial In Institutionalisierten Lehr-Lean-Prozessen-Am Beispiel Des Englischanfang-sunterrichts*. Gottingen: Deutsche Forschungsegemeinschaft, pp. 1753, n.p.

WILLIAMS, P. (1988). *A Glossary of Special Education*. Milton Keynes: Open University Press, pp. 222, £8·95 pbk., ISBN 0-335-15995-8.

WOLF, R. (1987). *International Journal of Educational Research Vol. 11, No. 1, The State of the Field*. Oxford: Pergamon Press, pp. 144, n.p.

ZARKOWSKA, E., and CLEMENTS, J. (1987). *Problem Behaviour in People with Severe Learning Disabilities*. Beckenham: Croom Helm, pp. 208, £10·95 pbk., ISBN 0-7099-3014-3.

## The Social Context of Emotion

A special issue of the *British Journal of Social Psychology*, guest edited by Hugh L. Wagner.

Introduction—*Hugh L. Wagner*
The social form of feeling—*Alan Radley*
Constructing emotion: Weaving meaning from memories—*Susan Kippax, June Crawford, Pam Benton, Una Gault & Jenny Noesjirwan*
Emotion/cognition interaction in personality development: A discrete emotions, functionalist analysis—*Carol Z. Malatesta & Arnold Wilson*
Emotion and social context: An American–German comparison—*Shula Sommers & Corinne Kosmitzki*
Sex differences in interpretation of angry and weeping episodes—*Muriel Egerton*
Impression management and self-deception: Response bias components of appraised emotional experience—*Gordon L. Flett, Kirk R. Blankstein, Patricia Pliner & Carl Bator*
Moods and compliance—*Sandra Milberg & Margaret S. Clark*

The special issue will be available as Part 1 of the *British Journal of Social Psychology*, volume 27 (1988). Subscription to volume 27 (ISSN 0144-6665), edited by Dr A. S. R. Manstead, is £49.50 (US$89.00)

In order to make it available to the widest possible audience, the special issue is available as a single part, for a limited period only (until June 1989) at a special price of £6.95 (US$11.95) for the single journal part.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

## Technological Change and Innovation

A special issue of the *Journal of Occupational Psychology*, guest edited by Chris Brotherton.

Strategic human resource management in the UK computer supply industry—*Paul R. Sparrow & Andrew M. Pettigrew*
A longitudinal study of adjustment to technological change: From mass to computer-automated batch production—*Ann Majchrzak & John Cotton*
Parental influence and teenagers' motivation to train for technological jobs—*Glynis M. Breakwell, Chris Fife Schaw & John Devereux*
The development of computer applications by non-employed people in community settings—*Catherine Cassell, Mike Fitter, David Fryer & Leigh Smith*
Long-wave cycles of sociotechnical change and innovation: A macropsychological perspective—*Kenyon B. De Greene*
Information technologies and organizations—lessons from the 1980s and issues for the 1990s—*Frank Blackler*
A history of social psychological reactions to new technology—*Jim Carliopo*
Robotics—A challenge for occupational psychology—*Oded Shenkar*

The special issue will be available as Part 1 of the *Journal of Occupational Psychology*, volume 61 (1988). Subscription to volume 61 (ISSN 0305-8107), edited by Dr David Guest, is £49.50 (US$89.00).

In order to make it available to the widest possible audience, the special issue is available as a single part, for a limited period only (until June 1989) at a special price of £6.95 (US$11.95) for the single journal part.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

## Clinical Attribution Research

A special issue of the *British Journal of Clinical Psychology*, guest edited by Chris R. Brewin

Editorial. Developments in an attributional approach to clinical psychology—*Chris R. Brewin*
The hopelessness theory of depression. Attributional aspects—*Lauren B Alloy, Lyn Y. Abramson, Gerald I Metalsky & Shirley Hartlage*
A longitudinal test of the attributional vulnerability model in children at risk for depression—*Constance Hammen, Cheri Adrian & Donald Hiroto*
Attributional change during psychotherapy—*Jenny Firth-Cozens & Chris R Brewin*
Attribution and adjustment to life-threatening illness—*Dawn C Turnquist, John H Harvey & Barbara Anderson*
Causal attributions of doctor and patients in a diabetic clinic—*C R. Gillespie & C Bradley*
The impact of attributions in marriage: Empirical and conceptual foundations—*Frank D. Fincham & Thomas N. Bradbury*
Causal beliefs amongst families in therapy: Attributions at the group level—*Anthony G Munton & Charles Antaki*
Attribution theory and attributional therapy: Some theoretical observations and suggestions—*Bernard Weiner*
Cognitive style and depression: Symptom-related, event-related or independent provoking factor?—*Glenys Parry & Chris Brewin*

The special issue will be available as Part 1 of the British Journal of Clinical Psychology, volume 27 (1988). Subscription to volume 27 (ISSN 0144-6657), edited by Professor Ray Hodgson, Whitchurch Hospital, Cardiff, is £49.50 (US$89.00).

Single copies of Part 1 only will be made available at the special price of £6.95 (US$11.95) for a limited period until June 1989 only.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

## Stress and Health

A special issue of the *British Journal of Medical Psychology*, guest edited by Cary L. Cooper and Sidney Crown.

Introduction—*Cary L. Cooper*
A structural model approach toward the development of a theory of the link between stress and mental health—*Cary L Cooper & A. J. Baglioni, Jr*
The impact of social support on mental and physical health—*Daniel C. Ganster & Victor Bart*
Type A behaviour and the healthy individual—*John M. Ivancevich & Michael T. Matteson*
Immunological consequences of acute and chronic stressors: Mediating role of interpersonal relationships—*Susan Kennedy, Janice K Kiecolt-Glaser & Ronald Glaser*
Environmental stress, reactivity and ischaemic heart disease—*David S. Krantz & Shera E Raisen*
Personality, stress and cancer: Prediction and prophylaxis—*H J. Eysenck*

The special issue will be available as Part 1 of the *British Journal of Medical Psychology*, volume 61 (1988). Subscription to volume 61 (ISSN 007-1129), edited by Dr Sidney Crown, is £59 00 (US$106.00).

In order to make it available to the widest possible audience, the special issue is available as a single part, for a limited period only (until June 1989) at a special price of £6.95 (US$11.95) for the single journal part.

Orders and subscriptions to:

**The British Psychological Society**
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

*Scottish Academic Press*

**PROFESSIONAL ISSUES IN EDUCATION Series**

Editors: Gordon Kirk and Robert Glaister

Contemporary issues of the Scottish education system are examined and discussed in this series. The contributors have been chosen for their specialised knowledge in particular areas of education.

**TEACHER EDUCATION AND PROFESSIONAL DEVELOPMENT VOL 1**

Gordon Kirk

There is a great deal of interest in current development in Scottish teacher education and these, along with other issues such as appraisal, conditions of service, professional responsibilities and the role of the teacher in the light of the Main Report, are discussed in this book.

*144 pages paperback £4.95 0 7073 0522 5*

**ASSESSMENT: A CHANGING PRACTICE VOL 2**

Sally Brown (editor)

Five interrelated themes on the issue of pupil assessment are addressed in this book. They are: assessment in relation to 'testing'; the range of qualities assessed; descriptive assessment; development of responsibilities for assessment and assessment for certification.

*152 pages paperback £4.95 0 7073 0515 2*

For further information please contact: **Sales Promotion Manager**, Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

# *Scottish Academic Press*

## PROFESSIONAL ISSUES IN EDUCATION Series

Editors: Gordon Kirk and Robert Glaister

Contemporary issues of the Scottish education system are examined and discussed in this series. The contributors have been chosen for their specialised knowledge in particular areas of education.

## PERSONAL AND SOCIAL EDUCATION VOL 3

John MacBeath

The issue of 'personal, social and vocational education' is tackled in this volume, and the three following questions are asked: How do the government and the public perceive education? What has been the role of researchers on education during the last decade? And how is education portrayed in media coverage?

*164 pages paperback £4.95 0 7073 0531 4*

## EDUCATION AND VOCATION 14-18 VOL 4

A Douglas Weir

The role of vocational education in the educational system as a whole is looked at in this book, with regard to the fact that unemployment is on the increase. It asks whether a higher priority is given to academic qualifications than to vocational skills and which is the most relevant in the economic and social climate of today.

*152 pages paperback £4.95 0 7073 0551 9*

For further information please contact: **Sales Promotion Manager**, Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

## PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

*British Journal of Psychology*
*British Journal of Medical Psychology*
*Journal of Occupational Psychology*
*British Journal of Social Psychology*
*British Journal of Clinical Psychology*
*British Journal of Developmental Psychology*
*British Journal of Mathematical and Statistical Psychology*
*The British Journal of Educational Psychology*

The *British Journal of Psychology*, the *British Journal of Medical Psychology, Journal of Occupational Psychology*, the *British Journal of Development Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1988 subscription prices, payable in advance, are £74·50 (US$134·00) for the *British Journal of Psychology;* £49·50 (US$89·00) for the *British Journal of Social Psychology,* £49·50 (US$89·00) for the *British Journal of Clinical Psychology;* £45·00 (US$81·00) for the *British Journal of Developmental Psychology,* £59·00 (US$106·00) for the *British Journal of Medical Psychology;* and £49·50 (US$89·00) for the *Journal of Occupational Psychology.*

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £49·50 per volume post free (US$89·00 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts., SG6 1HN, UK. A special price of £340·00 (US$£610·00) for all seven journals going to the same address is offered for 1988.

*The British Journal of Educational Psychology* is issued three times a year, in February, June and November. The prepaid subscription is £20·00 UK, £22·00 overseas a year, post free, or £7·00 each number. The prepaid subscription price in the USA is $45·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX

### MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D. E. Blackman, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Medical Psychology* should be sent to Dr. Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.

For the *Journal of Occupational Psychology* should be sent to Dr. D. Guest, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Social Psychology* should be sent to Dr. A. S. R. Manstead, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Clinical Psychology* should be sent to Professor R. Hodgson, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr. A D. Lovie, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Developmental Psychology* should be sent to Professor George Butterworth, Journals Office, 13A Church Lane, London, W2 8DX.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Hazel Francis, University of London Institute of Education, 24-27 Woburn Square, London, WC1H 0AA.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

**All manuscripts should be submitted in duplicate.**

### OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

### REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., PO Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

### SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £40·00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £7·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

# BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 58, PART 2, 1988

## CONTENTS

ARTICLES

DICKSON, G. C. A., and HUTCHISON, G. E. — Children's Perception of and Anticipated Responses to Risk ........ 147

CRASKE, M. L. — Learned Helplessness, Self-Worth Motivation and Attribution Retraining for Primary School Children ........ 152

BROOKS, M. R., GLENN, S. M., and CROZIER, W. R. — Pre-School Children's Preferences for Drawings of a Similar Complexity to Their Own ........ 165

HUNTING, R. P., and SHARPLEY, C. F. — Pre-Schoolers' Cognitions of Fractional Units ........ 172

HEWISON, J. — The Long-Term Effectiveness of Parental Involvement in Reading: a Follow-up to the Haringey Reading Project ........ 184

GLOVER, J. A., HARVEY, A. L., and CORKILL, A. J. — Remembering Written Instructions: Tab A Goes into Slot C, or Does It? ........ 191

RESEARCH NOTES:

WHITTINGTON, J. — Large Verbal — Non-Verbal Ability Differences and Underachievement ........ 205

COOK, J. — An Investigation of the Validity of the British Ability Scales with Respect to the Wechsler Intelligence Scale for Children — Revised and the Wide Range Achievement Test — Revised on a Group of Canadian Children ........ 212

MCKAY, M. F., and NEALE, M. D. — Patterns of Performance on the British Ability Scales for a Group of Children with Severe Reading Difficulties ........ 217

CONFERENCE ABSTRACTS ........ 223

BOOK REVIEWS ........ 231

PUBLICATIONS RECEIVED ........ 241

UK ISSN 0007-0998

# THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1988

VOL. 58, PART 3

# THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by
Professor HAZEL FRANCIS

Assistant Editors:
Dr. H. C. M. CARROLL
Prof. C. W. DESFORGES
Dr. D. FONTANA
Dr. F. J. TAYLOR

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Prof. C. W. DESFORGES

**Advised by the Editorial Committee:**

Professor G. BROWN
Ms. I. LUNT
Dr. C. LEACH
Professor P. W. ROBINSON
Mr. E. WILKINSON

Vol. 58, Part 3

*The British Journal of Educational Psychology* is issued by the British Psychological Society once a term, in February, June and November. The subscription is £20.00 a year to the UK, £22.00 overseas, or £7.00 each number.

*Subscriptions and Orders* should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

*Papers for publication* in duplicate should be sent to Professor Hazel Francis, University of London Institute of Education, 24-27 Woburn Square, London, WC1H 0AA. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to: Prof. C. W. Desforges, University of Exeter, St. Luke's, Heavitree Road, Exeter, EX1 2LU.

*Advertisements* for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

*All rights reserved*

© 1988. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY LINDSAY & CO LTD , EDINBURGH

*Br. J. educ. Psychol.*, 58, 247-257, 1988

# STUDY APPROACHES, PROCESSES AND STRATEGIES: ARE THREE PERSPECTIVES BETTER THAN ONE?

By CAROL SPETH and ROBERT BROWN
*(University of Nebraska-Lincoln, USA)*

SUMMARY. A sample of 383 students in educational psychology classes at a large American university completed three inventories, one based on cognitive psychology, one on approaches to learning, one on autonomous studying. A series of factor analyses of items and subscales showed common emphases that transcend differences in theoretical background and terminology: effort-intensive organised study; intention to understand, personalise and integrate the information being learned; strategic or context-sensitive study behaviour; and an unselective intention to reproduce information, impervious to external cues and without personal involvement.

## INTRODUCTION

RESEARCHERS have looked at university student learning from several theoretical perspectives. Some emphasise cognition, information processing, and memory; while others emphasise relationships of learning style, personality, motivation or intention, to study strategies, or to depth or kind of learning. A third perspective (summarised by Derry and Murphy, 1986), concerns training students to use specific learning strategies, but Rohwer (1984) invited researchers to look instead at the strategies students choose to employ on their own, either in routine studying or preparing for examinations.

The purpose of this study was to compare inventories from three theoretical perspectives: cognitive processes, approaches to learning, and autonomous studying, to identify similarities and differences in how they attempt to describe the most important facets of study activity, and see if one perspective might contribute to an understanding of the other two.

### *Learning Processes*

Working within the framework of cognition, information processing and memory rather than intention, motivation or personality, Schmeck (1983, p. 235) decided to "custom-build a measure of individual differences in the predisposition to use certain learning strategies". The Inventory of Learning Processes was intended to be a series of "behaviourally oriented statements" written in terms of the environment of the typical college student (Schmeck *et al.*, 1977). Schmeck built upon the idea of depth of processing in the memory proposed by Craik and Lockhart (1972).

### *Intention + Process = Approach*

A second coalition of researchers became convinced that learning consists of three related elements: intention, process, and outcome (Entwistle *et al.*, 1979). The central concept in this body of research is a distinction between two approaches to a learning task: one described as deep, meaning-oriented, transformational, or internalising; the other as surface-oriented, reproducing, or memorising. Marton and Säljö (1976a, 1976b) investigated what students learned when they were asked to read academic articles, and found that what they learned varied according to their approach to the task. Marton categorised students' responses as "deep" or "surface", also recalling Craik and Lockhart (1972). But Marton's categories, unlike Schmeck's, included the learner's intentions (Entwistle, 1987).

Some researchers (Entwistle and Wilson, 1970) within this coalition also identified a second difference among students, relating to their motivation to use the study methods or strategies which they already know or could acquire. Biggs (1978) reinforced the idea that students can take one of three paths to learning, each consisting of a motive and a strategy for studying. Factor analysis of his Study Processes Questionnaire yielded a utilising factor including fact-rote strategy, extrinsic motivation, and fear of failure; an internalising factor including meaning assimilation and intrinsic motivation; and an achieving factor including organised studying and need for achievement.

The Approaches to Studying Inventory (Entwistle and Ramsden, 1983) was a product of research on student learning, which included motivations and study methods as well as contrasting the intention to understand with the intention to reproduce the information being processed. This inventory also attempted to measure differences in style or kind of learning, as described by Pask (1976a) who identified two contrasting learning styles: holist and serialist; two strategies: comprehension learning and operation learning; and two pathologies: improvidence and globetrotting (Pask, 1976b). It also included the ideas of "syllabus-boundness" described by Hudson (1968) and Parlett (1970), "cue-deafness" suggested by Miller and Parlett (1974), and "strategic approach" based on responding to contextual cues on what study behaviours are most rewarded by the assessment system (Entwistle *et al.*, 1979).

### *Autonomous Academic Studying*

Finally, J. W. Thomas and Rohwer (1986) described a model of "autonomous learning" with two broad categories of study activities. Cognitive transformational activities include selection, comprehension, memory, integration, and cognitive monitoring. Self-management activities include time management, effort management and volitional monitoring. Two instruments based upon this model were used to compare study activities in different subject areas in junior high, senior high, and college (Strage *et al.*, 1987). In some ways this perspective is similar to that of Schmeck (1983), but lacks the emphasis on individual differences which he shares with researchers in the approaches to learning coalition.

### *Combining perspectives*

Suggesting that the best and most useful inventories will eventually come from combining different theoretical bases, contexts, and item pools, Entwistle and Waterston (1985) administered and factor analysed subsets of approach to studying and learning processes items and subscales together. This study differs from theirs in sampling from an American rather than British university student population, using a slightly different version of the Approaches to Studying Inventory, and including subscales from the autonomous learning perspective.

## METHOD

### *Instruments*

The Approaches to Studying Inventory (ASI) consists of 64 statements about how students tackle everyday learning tasks; there are 16 subscales with four items each. The 16 subscales are: Deep Approach, Use of Evidence, Relating Ideas, Intrinsic Motivation, Surface Approach, Syllabus-Boundness, Fear of Failure, Improvidence, Extrinsic Motivation, Strategic Approach, Achievement Motivation, Disorganised Study Methods, Negative Attitudes, Globetrotting, Comprehension

Learning and Operation Learning. The 16 subscales are combined into four scales, each representing an orientation or approach to studying. Reliabilities obtained in this study for the four scales (Meaning, Achieving, Reproducing and Non-Academic) ranged from 0·58 to 0·73, similar to those reported by Ramsden (1983) and Entwistle and Waterston (1985). A few changes in wording were made to fit these items into an American context. A five-point Likert scale was used to conform with the standard mark-sense sheets.

The Inventory of Learning Processes (ILP) was originally made up of 62 statements written by experts in human learning and memory to describe typical study behaviours of college students (Schmeck *et al.*, 1977). The four subscales were called: Deep Processing, Elaborative Processing, Fact Retention, and Methodical Study. For this study, the same wording and reverse scoring recommended by Schmeck were used; however, a five-point Likert scale was substituted for true-false, and the number of items was reduced to 30, based on Entwistle and Waterston (1985). The four scales had Cronbach's alphas ranging from 0·67 to 0·78.

A Test Preparation Activities Survey was developed by adapting 55 items appropriate for the context of a college-level history survey course from the Study Activity Survey Form T (a 167-item questionnaire from the Far West Laboratory for Educational Research and Development in San Francisco). Principal axis factor analysis of the 55 items, followed by varimax rotation, was used to identify groups of related items. An eight-factor solution was chosen based on the scree test. These eight factors accounted for 46 per cent of the total variance. Two negligible factors were dropped, yielding six subscales. The six subscales, with their Cronbach's alpha coefficients of reliability were called: Time-Effort, 20 items, 0·87; Integration, 13 items, 0·78; Selection, 5 items, 0·79; Cognitive Monitoring, 5 items, 0·67; Peer Orientation, 5 items, 0·69; Average Student Behaviours, 7 items, 0·52.

The names of test preparation subscales came from the research literature and approximated item content as closely as possible. However, because these subscales were empirical rather than conceptual, it would be wrong to assume an ideal match between item content and the name of each factor. These groupings of items were fairly consistent across factor analyses using different criteria. Time-Effort could be defined as ensuring adequate time for studying, minimising competing demands, ensuring adequate attention and effort. Integration could be defined as constructing relationships among items to be learned, and between items to be learned and other information. Strage *et al.* (1987), who used all 167 items, distinguished between time and effort, but that distinction did not seem warranted on the basis of factor analyses of data from this sample with this subset of items. Selection could be defined as identifying cues, differentiating and prioritising; Cognitive Monitoring as awareness and control of one's cognitive resources in relationship to task demands (J. W. Thomas and Rohwer, 1986). Peer Orientation reflected the extent to which students rely on each other for advice on how to study, as opposed to relying on authority or contextual cues; to some extent this subscale was an index of social motivation. Reliability of the Average Student subscale was modest, but it was maintained because this factor emerged consistently from different types of factor analyses specifying different criteria. One discernible common element was that all the items on that factor seemed to describe behaviours typical of the average college student, study strategies representing a rather low-surface approach. Perhaps the closest concept in the research literature would be that of "skimming", described by P. R. Thomas and Bain (1984) as a filtering process relatively uncorrelated to either transformational (integrating, elaborating) or reproductive (selecting, remembering) activities. Scores for each student on each subscale were calculated and entered into a data file along with subscale scores on the other two instruments.

### *Sample*

Student subjects were recruited in educational psychology classes at a large state university in the United States. The 383 students included 125 male, 258 female. Approximately 84 per cent were between 18 and 25.

## RESULTS

Fundamental questions about the structure of each inventory would be addressed by a series of SPSSX principal axis factor analyses of (a) the 64 ASI items; (b) the 30-item ILP used in this study; (c) 45 ASI with 30 ILP items; and (d) the test preparation activities survey. The factor analysis combining items from the ASI and ILP was meant to identify their common and their unique characteristics: specifically, motivation on the ASI and memory on the ILP. A series of SPSSPC+ principal axis factor analyses of the subscales was undertaken to identify similarities between the psychologies of autonomous learning, approaches to studying, and cognitive processing.

### *Factor analyses of items*

The ASI items were submitted to principal axis factor analysis followed by varimax rotation. An eight-factor solution was chosen, based on the scree plot. These eight factors accounted for 37 per cent of the total variance. Factor I clearly represented Meaning Orientation. One Globetrotting item crept in with the four Comprehension Learning items. Factor II was dominated by disorganised study items. Factor III seemed to be Surface Approach with items phrased in different directions. Factor IV was the motivation factor, with extrinsic positive and intrinsic negative. Factor V included the negative attitudes and motivation characteristic of Non-Academic Orientation. Factor VI included academic behaviours typical of the Strategic Orientation; while Factor VII items were more typical of the Reproducing Orientation. Factor VIII only had two items, both on Achievement Motivation.

When the 30 ILP items were submitted to principal axis factor analysis, eight factors had eigenvalues above unity, and accounted for 58 per cent of the total variance. Factor I included both Deep-to-Surface and Elaborative Processing items. Factor II emphasised Fact Retention and memory. Nearly all items on Factor III came from the Methodical Study scale. Fact Retention items loaded not only on Factor II, but also on IV, V, and VII, and Elaborative Processing items appeared on Factors I, VI and VIII.

The same 30 ILP items can also be considered in combination with a subset 45 of ASI items, imitating Entwistle and Waterston (1985), but using principal axis rather than principal components. The eight-factor solution accounted for 42 per cent of the total variance. Factor I represented the Reproducing Orientation, with eight Deep-to-Surface and two Fact Retention items from the ILP. These items concerned having difficulty with: essay tests, learning how to study for each course, remembering, comparing concepts, making inferences, organising, critical evaluation and finding the right words. Seven ASI items also loaded on this factor, including fear of failure, test anxiety, tension and depression, getting distracted easily, feeling a need to memorise to survive, introducing irrelevant material into essays, and inability to see the overall picture. Factor II included five ILP and three ASI items, and represented the Non-Academic Orientation with a combination of low motivation and disorganised study methods. Factor III had four ILP and two ASI items and portrayed students who do not do a lot of thinking. Factor III included items such as "I rarely think", or "I rarely look for reasons", and "I usually do not think about implications". Factor IV, with seven items, four items from the Negative Attitudes subscale, concerned academic Achievement Motivation and its absence. Factor V had five Elaborative Processing items from the ILP and a

Deep Approach item from the ASI, and included visualising, problem-solving, summarising and putting the material to be learned into one's own words. Five Fact Retention items loaded on Factor VI. Three ASI items loaded on Factor VII, which concerned vocational or Extrinsic Motivation. Four ASI items on Factor VIII described a competitive, achievement-oriented student who jumps to conclusions and hates to admit defeat.

### *Factor analyses of subscales*

Factor analyses of subscales rather than items produced results which were easier to interpret and accounted for more variance. When the 16 ASI subscales were subjected to principal axis factor analysis followed by varimax rotation, four factors had eigenvalues above 1·0 and accounted for 53 per cent of the total variance. As Table 1 indicates, Factor I appears to be a combination of the reproducing and non-academic subscales. Factor II is made up of meaning-oriented subscales; Factor III includes Strategic Approach, Operation Learning, and Syllabus Bound. Factor IV is confined to motivation, both intrinsic and extrinsic. The highest loading for Intrinsic Motivation was on Factor IV, but it also loaded at 0·42 on the Meaning Orientation Factor. Syllabus Bound also loaded with Surface Approach at 0·33.

TABLE 1

ROTATED FACTOR MATRIX FOR THE 16 ASI SUBSCALES
LOADINGS ABOVE 0·30 OR HIGHEST IF BELOW 0·30

| | I | II | III | IV |
|---|---|---|---|---|
| DEEPAPRH | | 0·65 | | |
| RELATID | | 0·65 | | |
| USEVIDNC | | 0·54 | | |
| INTRNSIC | | 0·42 | | 0·73 |
| CMPLEARN | | 0·40 | | |
| ACHIEVMT | | | 0·22 | |
| EXTRNSIC | | | | −0·34 |
| STRATEGC | | | 0·47 | |
| OPRLEARN | | | 0·62 | |
| SURFACEA | 0·57 | | | |
| SYLBOUND | 0·33 | | 0·53 | |
| FEARFAIL | 0·54 | | | |
| IMPROVID | 0·50 | | | |
| DISORGSM | 0·51 | | | |
| NEGATIVE | 0·51 | | | |
| GLOBTROT | 0·60 | | | |

With the addition of the four ILP scales, five factors had eigenvalues above unity and accounted for 56 per cent of the total variance. Table 2 shows that Factor I had Deep-to-Surface loading at −0·77, Elaborative Processing at −0·30 and Fact Retention at −0·46; positive loadings included Surface Approach, Fear of Failure, Improvidence, Disorganised Study Methods, Negative Attitudes and Globetrotting. Factor II combined Elaborative Processing with the meaning-oriented subscales. Factor III represented Strategic Approach, but once again, included Syllabus Bound, as if to say that, in this context, being strategic requires obeying the "letter of the law" as far as assignments are concerned. Factor IV represented a separate study methods factor, Factor V represented motivation. The motivational component of the ASI tended to come out as a separate factor, while the memory component of the ILP did not.

TABLE 2

Rotated Factor Matrix for the 20 ILP and ASI Subscales
Loadings above 0·30 or highest if below 0·30

| | I | II | III | IV | V |
|---|---|---|---|---|---|
| DEEPSURF | −0·77 | | | | |
| ELABPROC | −0·30 | 0·64 | | | |
| STDYMETH | | | | 0·92 | |
| FACTRETN | −0·46 | | | | |
| DEEPAPRH | | 0·63 | | | |
| RELATID | | 0·61 | | | |
| USEVIDNC | | 0·54 | | | |
| INTRNSIC | | 0·46 | | | 0·64 |
| CMPLEARN | | 0·43 | | | |
| ACHIEVMT | | | 0·22 | | |
| EXTRNSIC | | | 0·30 | | −0·25 |
| STRATEGC | | | 0·41 | | |
| OPRLEARN | | | 0·63 | | |
| SURFACEA | 0·49 | | 0·37 | | |
| SYLBOUND | | | 0·60 | | |
| FEARFAIL | 0·53 | | | | |
| IMPROVID | 0·51 | | 0·31 | | |
| DISORGSM | 0·39 | | | −0·59 | |
| NEGATIVE | 0·43 | | | | |
| GLOBTROT | 0·62 | | | | |

When the six test preparation subscales were added, six factors had eigenvalues above unity and accounted for 57 per cent of the total variance. Table 3 indicates that Factors I and II from the previous analysis traded places. Intrinsic Motivation and Integration loaded with the meaning-oriented subscales. Factor II includes loadings of −0·68 for Deep-to-Surface, −0·71 for Fact Retention, and −0·39 for Cognitive Monitoring. Selection (which could have been called "cue-deafness") loaded strongly on this factor in a positive direction (0·69), along with Globetrotting (0·66) and Negative Attitudes (0·45). Subscales for the Reproducing Orientation, plus Operation Learning, loaded on Factor IV. Factor III included Study Methods at −0·86 and Disorganised Study Methods at 0·80, along with Time-Effort at 0·85. Factor V included Peer Orientation and Average Student. Factor VI represented Strategic Orientation, and included a 0·64 loading for Achievement Motivation, 0·47 for Extrinsic and 0·38 for Strategic.

## DISCUSSION

Resemblance between individual factors and one of the four approaches was complicated by the eight-factor solution, but can be seen nevertheless. Meaning Orientation solidified on Factor I, while Reproducing loaded on II and VII. Non-Academic solidified fairly well on Factor V. But Strategic Orientation split apart,

TABLE 3

ROTATED FACTOR LOADINGS ABOVE 0·30 FOR ILP, ASI AND TEST PREPARATION SUBSCALES

| | I | II | III | IV | V | VI |
|---|---|---|---|---|---|---|
| DEEPSURF | | −0·68 | | −0·47 | | |
| ELABPROC | 0·51 | | | −0·48 | | −0·31 |
| STDYMETH | | | −0·86 | | | |
| FACTRETN | | −0·71 | | | | |
| DEEPAPRH | 0·71 | | | | | |
| RELATID | 0·68 | | | | | |
| USEVIDNC | 0·68 | | | | | |
| INTRNSIC | 0·65 | | | | | −0·31 |
| CMPLEARN | 0·51 | | | −0·35 | | |
| ACHIEVMT | | | | | | 0·64 |
| EXTRNSIC | −0·30 | | | | | 0·47 |
| STRATEGC | | | | | | 0·38 |
| OPRLEARN | | | | 0·68 | | 0·30 |
| SURFACEA | | 0·33 | | 0·51 | 0·32 | |
| SYLBOUND | | | | 0·63 | | |
| FEARFAIL | | 0·46 | | 0·48 | | |
| IMPROVID | | | | 0·58 | | |
| DISORGSM | | | 0·80 | | | |
| NEGATIVE | | 0·45 | | | | |
| GLOBTROT | | 0·66 | | | | |
| TIMEFFRT | | | 0·85 | | | |
| INTEGRAT | 0·52 | | −0·49 | | 0·37 | |
| SELECTN | | 0·69 | | | | |
| COGMONIT | 0·31 | −0·39 | | | | |
| PEERORI | | | | | 0·77 | |
| AVERSTU | | | | | 0·77 | |

loading on Factors II, IV, VI and VIII, which makes sense because an essential characteristic of that approach is being able to use either deep or surface strategies depending on context.

The results of factor analysing the 30 ILP items were not what one would expect based on Schmeck *et al.* (1977) who found four factors with eigenvalues above unity when they analysed all 62 items. Their first factor was clearly Synthesis-Analysis (later called Deep Processing), while Study Methods emerged second, Fact Retention third, and Elaborative Processing fourth. In this study, Factor II emphasised Fact Retention, and Factor III clearly emphasised Study Methods, but the other factors were not very clear. While the scale reliabilities for the shortened ILP were quite comparable to the 62-item version, cutting the number of items and changing the response format from true-false to a Likert scale may have hurt the factor analysis. However, these results made more sense than those reported by Watkins and Hattie (1981) who found that factor analyses of the 62 ILP items for their Australian sample "did not produce any solution which could be interpreted in terms of the present ILP scales".

It is possible to compare results obtained by this factor analysis of 30 ILP items with a subset of 45 ASI items with the six factors Entwistle and Waterston (1985) identified, although there were differences in the selection of items and the direction in which some of them were phrased. In their study, Schmeck's "Deep Processing" scale was called "Surface Processing", with changes in the direction of some items; while in this study, Schmeck's items were used verbatim, and that scale was called "Deep-to-Surface". Their first factor, like Factor I in this study, brought together surface processing with Reproducing Orientation, learning pathologies and operation learning, and they said that "the intention to reproduce is linked to low levels of critical evaluation and conceptual organisation". They described Surface Approach in terms of failure to think analytically. Their second factor, similar to, yet more comprehensive than Factor V in this study, suggested an overlap between Elaborative Processing, Meaning Orientation and positive attitudes to studying, "thus the intention to understand seems to involve personalising the information presented and relating it to real-life situations". Their third factor was most like Factor II, with Disorganised Study Methods and Negative Attitudes. Their fourth factor was much like Factors IV and VIII, and they said it "describes a self-confident, competitive and strategic approach to studying". Their analysis also separated the organised study from the strategic factors.

In this study, when ASI items were considered, Non-Academic behaved like a separate orientation, not just the low end of Surface Approach — the impression given by factor analysis of 16 or 20 subscales (Tables 1 and 2). When 26 subscales were considered (Table 3), Factor I included Elaborative Processing plus Meaning Orientation. Factor II which included Deep-to-Surface (−0·68), Fact Retention (−0·71), Cognitive Monitoring (−0·31), Selection (0·69), Negative Attitudes (0·45) and Globetrotting (0·66), might be considered Non-Academic + low Surface Approach. Factor IV, with Surface Approach (0·51), Syllabus Bound (0·63), Fear of Failure (0·48), and Improvidence (0·58), might be considered high Surface Approach, closer to Strategic.

On the last factor analysis of 26 subscales, Intrinsic Motivation moved into the meaning-oriented factor, suggesting that with a larger number of more reliable elements from the same domain, and more subjects per element, the motivation subscales might line up where they are supposed to, rather than on a separate factor (Tables 1-2), a result which may be peculiar to this study. Factor V on Table 3 suggests that for this sample of American students Average Student and Peer Orientation were more associated with Reproducing than Non-Academic Orientation. Entwistle and Waterston (1985) found that their social motivation items loaded on the non-academic factor. It may be that students considered non-academic in Britain would be considered surface-oriented at a midwestern American state university because of different admissions standards.

Previous studies (Entwistle and Ramsden, 1983) have consistently identified deep and surface approach factors, and they were clearly identified here as well. A third factor representing strategic or achievement motivation has also appeared consistently in previous studies, but with small variations in composition. More than deep *vs*. surface approach, one would expect motivational patterns to be influenced by characteristics of specific learning environments and by cultural, educational, and economic conditions in various countries at different times. Although the defining features of each approach are consistent, for supplementary characteristics (like Syllabus Bound) some shifting around on factors may be inevitable. This implies there may not be one best way to combine ASI subscales which will cross international boundaries and fit all educational contexts (for more discussion, see Speth, 1987, pp. 111-112).

In this study, factor analyses combining two or more perspectives generally supported the suggestion by Entwistle and Waterston (1985, p. 8) that "we seem justified in being able to describe study approaches and processes in terms of four dimensions: deep/elaborative, surface, organised, and strategic/competitive". The autonomous studying perspective contributes an element of time- and effort-intensity to the organised dimension and reinforces the cue-deafness component of the surface dimension. The test preparation activities survey developed for this study included only certain parts of the model of academic studying proposed by Thomas and Rohwer (1986). Further research in this area could strive for a more complete representation of their ideas in order to gain a better sense of the linkages with the other two perspectives. If more memory items had been included, for example, more could be said about the Fact Retention aspect of the ILP. Revising and improving the Cognitive Monitoring subscale might clarify its relationship to the deep/elaborative dimension. For some reason, there seemed to be somewhat more unexplained variance with the American than with the British student samples, but analysing subscales rather than items, and adding subscales developed from other research bases led to gradual improvements in explained variance (comparable to those reported by Ramsden and Entwistle, 1981), suggesting that three perspectives might have more explanatory power than one or two. Of the three perspectives, the approaches to studying seemed to furnish the best interpretive framework for understanding the other two. At least four common emphases were identified, transcending the differences in the inventories' theoretical bases and terminologies: time- and effort-intensive organised study methods; an intention to understand, personalise and integrate the information being learned; strategic or context-sensitive study behaviour; and an unselective intention to reproduce information, impervious to external cues and without personal involvement.

ACKNOWLEDGMENTS.—Drs. R. R. Schmeck, N. J. Entwistle, P. Ramsden and J. W. Thomas gave permission to adapt their inventories for this study; Dr. Entwistle gave valuable suggestions at several stages of the research.

Correspondence and requests for reprints should be addressed to Carol Speth, Department of Educational Psychology, University of Nebraska-Lincoln, Lincoln, Nebraska, USA 68588-0641.

## REFERENCES

BIGGS, J. B. (1978). Individual and group differences in study processes. *Br. J. educ. Psychol.*, 48, 266-279.

CRAIK, F. I., and LOCKHART, R. S. (1972). Levels of processing: a framework for memory research. *J. verb. Learn. verb. Behav.*, 11, 671-684.

DERRY, S. J., and MURPHY, D. A. (1986). Designing systems that train learning ability: from theory to practice. *Rev. educ. Res.*, 56, 1-39.

ENTWISTLE, N. J. (1987). Generative concepts and pedagogical fertility: communicating research findings on student learning. (Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C.)

ENTWISTLE, N. J., HANLEY, M., and HOUNSELL, D. (1979). Identifying distinctive approaches to studying. *Higher Educ.*, 8, 365-380.

ENTWISTLE, N. J., and RAMSDEN, P. (1983). *Understanding Student Learning*. London: Croom Helm.

ENTWISTLE, N. J., and WATERSTON, S. (1985). Approaches to studying and levels of processing: a comparison of inventories derived from different theoretical bases. (Paper presented at International Conference on Cognitive Processes in Student Learning, Lancaster University, July 18-21, 1985.)

ENTWISTLE, N. J., and WILSON, J. D. (1970). Personality, study methods and academic performance. *Univ. Q.*, 24, 147-156.

HUDSON, L. (1968). *Frames of Mind*. London: Methuen.

MARTON, F., and SALJÖ, R. (1976a). On qualitative differences in learning: I. Outcome and process. *Br. J. educ. Psychol.*, 46, 4-11.

MARTON, F., and SALJÖ, R. (1976b). On qualitative differences in learning. II. Outcome as a function of the learner's conception of the task. Symposium: Learning processes and strategies. *Br. J. educ. Psychol.*, 46, 115-127.

## APPENDIX

### INDICATIVE INVENTORY ITEMS FOR SUBSCALES OF THE TEST PREPARATION ACTIVITIES SURVEY

*Time-Effort*

- When studying for this test outside of class, I would probably be able to plan my time well for each study session.
- When studying for this test outside of class, I would probably have trouble sticking to the task.

*Integration*

- If I were studying for this test outside of class, I would probably make a chart or picture of how different topics fit together.
- During the last class meeting before the test, I would probably think about how points raised in class are related to ideas and facts I have learned outside of this course.

*Selection*

- If I were studying for this test outside of class, I would probably have trouble deciding what was important.
- If I were studying for this test outside of class, I would probably have trouble figuring out what would be on the test.

*Cognitive Monitoring*

- When studying for this multiple choice test outside of class, I would probably be able to improve my understanding when I came across something that was hard to understand.
- When studying for this multiple choice test outside of class, I would probably be able to make it more memorable when I came across something important that was hard to remember.

*Peer Orientation*

- When I came across something important that I did not understand, I would probably try to get an explanation from another student.
- When something presented in the last class meeting before the test is hard to understand, I would probably ask another student for an explanation.

*Average Student*

- If I were studying for this multiple choice test outside of class, I would probably skim through the study materials to figure out how much work had to be done.
- When I came across something which I thought would be hard to remember, I would probably try to learn it by reading and rereading.

*Br. J. educ. Psychol.*, 58, 258-265, 1988

# APPROACHES TO STUDYING AND LEVELS OF PROCESSING IN UNIVERSITY STUDENTS

By NOEL ENTWISTLE AND SUSAN WATERSTON
*(University of Edinburgh)*

SUMMARY. Attempts at measuring students' study styles and strategies have been derived from different theoretical standpoints. Traditionally these have emerged from a combination of mainstream psychology and personal experience. More recently concepts have been derived from the qualitative analysis of students' experiences of studying and subsequently operationalised through inventories. This study compares the dimensions emerging from these two contrasting approaches. A sample of 218 mainly first-year university students were given shortened versions of two inventories, one based on concepts from cognitive psychology and the other from educational research. Correlations between subscales from the two inventories, and factor analyses of them, are reported which show substantial agreement between the main concepts used from the two theoretical traditions to describe study strategies.

## INTRODUCTION

THIS study was designed to compare two inventories, derived from contrasting theoretical bases, which have been developed to describe students' learning approaches and processes.

Previous attempts at describing important facets of studying concentrated initially on atheoretical approaches which catalogued those activities considered to be part of "good" study methods or to indicate high levels of academic motivation (Brown and Holtzmann, 1966; Entwistle and Wilson, 1970). Such attempts are continuing with increasing sophistication and seem likely to develop into extensive coherent descriptive frameworks, from which theoretical rationales should emerge (Rohwer, 1984; Thomas, 1985). But other work has attempted to develop inventories which are derived from pre-existing theoretical rationales.

The burgeoning research in the field of cognitive psychology provides one way of rooting study processes in theory. Schmeck (Schmeck *et al.*, 1977) reported the development of a 62-item Inventory of Learning Processes which was derived from the factor analysis of a pool of 121 self-report items written by a team of specialists within the research domain of human learning. They identified from the research evidence important learning processes and converted these into behavioural indicators related to the context of everyday studying. The underlying framework was provided by the notion of *levels of processing* in memory introduced by Craik and Lockhart (1972). Evidence for the validity of the inventory can be derived from the wide range of studies reported by Schmeck (1983).

An alternative theoretical framework is provided by research reported in *The Experience of Learning* (Marton *et al.*, 1984). Instead of anticipating the dimensions of study processes from experimental work in cognitive psychology, it is possible to derive important descriptive concepts and categories from qualitative analyses of students' reports of their own study processes. The main concept emerging from this research describes *approaches to learning*. These approaches contain a similar distinction between deep and surface levels of processing as that investigated by Schmeck, but Marton characterises the distinction between these categories in terms of intentionality (Marton and Säljö, 1984). A deep approach involves an intention to understand, while a surface approach describes the intention to reproduce information in compliance with externally imposed task demands. A third approach was introduced by Ramsden (1981). He described a *strategic approach* through which students seek to maximise their grades by strategic management of their time and intellectual resources in line with the perceived criteria for high grades.

These three approaches formed the basis of a 64-item Approaches to Studying Inventory (Entwistle *et al.*, 1979; Entwistle and Ramsden, 1983). Items were derived mainly from the interview comments of students which had formed the defining features of the three approaches found in the qualitative analyses. In addition, items distinguishing between Pask's (1976) *operation* and *comprehension* learning styles, and his *improvidence* and *globetrotting* pathologies were included. A final group of subscales were derived from analyses of an earlier inventory of motivation and study methods (Entwistle, 1975). Evidence for the reliability and validity of the inventory has been discussed elsewhere (Entwistle and Ramsden, 1983; Entwistle, 1988; Watkins, 1983).

## METHOD

### *Development of the Inventory*

The substantial length of the two inventories combined was considered likely to reduce the probability of completion of this voluntary task. It was therefore decided to use previously reported item analyses (Schmeck *et al.*, 1977; Entwistle and Ramsden, 1983) to select the 30 items from each inventory which provided the best combination of coverage and discrimination. A further group of 15 items were included to cover the further defining features of the main constructs which had emerged since the original scales had been devised, to help to explore areas of overlap between the inventories, and to create one additional scale (social motivation) suggested by the work of Taylor (1983). The final inventory consisted of 75 items.

Previous factor analyses of the Approaches to Studying Inventory (Entwistle and Ramsden, 1983) had suggested that there were four main orientations to studying into which approaches to learning, styles of learning, and associated forms of motivation merged (Entwistle, 1988). These four orientations were described as *achieving* (with components covering strategic approach, hope for success, and vocational motivation), *meaning* (deep approach and intrinsic motivation), *reproducing* (surface approach and fear of failure) and *non-academic* (disorganised study methods, negative attitudes, and social motivation). Subscales based on Pask's pathologies and on his serialist (operation learning) style were also included.

The Inventory of Learning Processes has consistently yielded the four main factors described by Schmeck (1983). *Deep processing* describes the extent to which a student critically evaluates, conceptually organises, and compares and contrasts information being studied. *Elaborative processing* indicates the extent to which students translate new information into their own terminology, apply it to their own lives, generate concrete examples from their own experience, and use visual imagery for the purpose of encoding new information. *Fact retention* involves perceived facility at learning facts and details, while *methodical study* contains activities recommended by "How to Study" manuals.

As the original formats of the inventories differed, response and scoring procedures were standardised as a five-point Likert scale with items scored all in the same direction. Schmeck had used a two-point Likert scale and items which varied in their direction of scoring. It is, of course, necessary to vary the direction of wording in the inventory as a whole to avoid an acquiescence response set, but this can be achieved equally well by including complete scales defined in both "positive" and "negative" directions (in relation to what students would perceive to be "good" study methods). Thus Schmeck's scales were reworded so that all the items in each scale were scored in the same direction, but two of his scales were changed to have a "negative" direction. "Deep processing" was changed to *surface processing* and "methodical" to *disorganised study habits.* In this way the acquiescence response set was again avoided.

### *Sample*

The shortened combined inventory was administered to volunteers, mainly first-year students at Edinburgh University, towards the end of the winter term. There were 117 science students (93 men and 24 women) and 101 arts and social science students (29 men and 72 women).

## RESULTS

Confirmatory factor analysis was initially carried out separately on the items within each part of the inventory to ensure that the scoring procedure could be justified in relation to the factor structure. The items within the short versions of the various subscales fitted the pattern found in the longer versions very closely. Next the internal reliability of the main subscales was investigated by calculating Cronbach "alpha" coefficients. Six of the scales showed a reasonably high level of internal consistency (r > 0·7). Lower consistency was found for the achieving orientation subscale (attributable to two items which represented a rather different dimension than the other items), while the consistency of the non-academic

TABLE 1

INTER-CORRELATIONS BETWEEN SCHMECK AND ENTWISTLE SUBSCALES

| | SCHMECK SUBSCALES | | | |
|---|---|---|---|---|
| | Surface Proc. | Disorg. Study | Fact Retention | Elab. Proc. |
| Entwistle Subscales | | | | |
| *Achieving Orientation* | | | | |
| Strategic Approach | | −0·12 | 0·15 | 0·20 |
| Hope for Success | | | 0·16 | 0·19 |
| Vocational Motivation | | | | |
| *Meaning Orientation* | | | | |
| Deep Approach | −0·26 | −0·28 | | 0·64 |
| Intrinsic Motivation | 0·30 | 0·32 | | 0·50 |
| *Reproducing Orientation* | | | | |
| Surface Approach | 0·50 | 0·18 | | −0·17 |
| Fear of Failure | 0·47 | | | |
| *Non-Academic Orientation* | | | | |
| Disorganised Approach | 0·29 | 0·67 | −0·14 | −0·14 |
| Negative Attitudes | 0·34 | 0·34 | −0·10 | −0·30 |
| Social Motivation | | 0·37 | | |
| *Pathologies and Styles* | | | | |
| Globetrotting | 0·37 | 0·18 | −0·11 | −0·17 |
| Improvidence | 0·51 | 0·15 | | −0·15 |
| Serialist Style | 0·36 | | | |

Correlations less than 0·10 omitted

orientation scales was similarly weakened by two of the social motivation items. The attempts to measure Pask's dimensions in this shortened version of the inventory were not successful, judging by the internal reliabilities of below 0·5. Nevertheless these scales were retained in the subsequent investigation of relationships between subscales, in case any indicative relationships might still be found.

To explore the inter-relationships between the two inventories it was decided to utilise the component parts of the subscales derived from the Approaches to Studying Inventory. Although the reliability of these components will be somewhat lower than those reported above, these groups of items are conceptually pure and it was considered that their use would facilitate interpretation of the relationships with the levels of processing subscales. Table 1 presents the statistically significant intercorrelations between the four subscales of learning processes and the 13 components of approaches to studying.

Surface Processing shows its highest correlations with Surface Approach, Fear of Failure, and Improvidence. Elaborative Processing is closely related to the Deep

TABLE 2

FACTOR STRUCTURE FOR TOTAL SAMPLE AND FOR SCIENCE AND ARTS/SOCIAL SCIENCE SAMPLES

| Subscales and Components | Total Sample (N = 218) | | | | Science (N = 117) | | | | | Arts/Social Science (N = 101) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I | II | III | IV | I | II | III | IV | V | I | II | III | IV | V |
| *Schmeck Scales* | | | | | | | | | | | | | | |
| Surface Processing | 73 | | | | 74 | | | | | 61 | | | | |
| Disorganised Study Methods | | | 83 | | | −41 | 85 | | | | | 85 | | |
| Fact Retention | | | | 61 | −45 | | | 53 | | | | | 81 | |
| Elaborative Processing | | 81 | | | | 87 | | | | | 82 | | | |
| *Achieving Orientation* | | | | | | | | | | | | | | |
| Strategic Approach | | | | 61 | | | | 75 | | 50 | 50 | | | |
| Hope for Success | | | | 45 | | | | | 83 | | | | 74 | −41 |
| Vocational Motivation | | | | 57 | | | | 55 | | 55 | | | | |
| *Meaning Orientation* | | | | | | | | | | | | | | |
| Deep Approach | | 83 | | | | 89 | | | | | 85 | | | |
| Intrinsic Motivation | | 79 | | | | 65 | | | 56 | | 78 | | | |
| *Reproducing Orientation* | | | | | | | | | | | | | | |
| Surface Approach | 77 | | | | 79 | | | | | 70 | −40 | | | |
| Fear of Failure | 76 | | | | 71 | | | | | 77 | | | | |
| *Non-Academic Orientation* | | | | | | | | | | | | | | |
| Disorganised Approach | | | 82 | | | | 82 | | | | | 85 | | |
| Negative Attitudes | | −46 | (38) | | 41 | (−35) | 50 | | | | −52 | | | |
| Social Motivation | | | 67 | | | | 65 | | | | | 54 | | |
| *Pathologies & Style* | | | | | | | | | | | | | | |
| Globetrotting | 52 | | | | 56 | | | | | | | | | |
| Improvidence | 69 | | | | 58 | | | | | 73 | | | | |
| Serialist | 66 | | | | 66 | | | | | 74 | | | | |

Decimal points, and loadings less than 0·40, omitted

Approach and Intrinsic Motivation. Disorganised Study Habits is very closely correlated with Disorganised Approach. However, Fact Retention shows only weak relationships with the components of approaches to studying, while the Achieving Orientation components are similarly unconnected with levels of processing.

The SCSS program (Nie *et al.*, 1975) was used to carry out principle component factor analysis with rotation to oblique simple structure (with delta set at zero). This procedure was carried out first for the total sample, and then separately for the science and arts/social science samples. The criterion of eigen values above unity was used to extract four factors in the total sample and five factors in both of the other samples. The factor structure matrices are shown in Table 2. The percentages of the variance explained by the factors were 55 per cent (overall), 62 per cent (scientists), and 64 per cent (arts/social sciences). The intercorrelations between the factors were uniformly low (none exceeding 0·17) and are thus not reported here.

The factor structures show a very clear pattern linking the two parts of the inventory in the ways already anticipated from the simple correlations. The first factor brings together Surface Processing with Reproducing Orientation, Pathologies and Serialist Style. Thus the intention to reproduce is linked to low levels of critical evaluation and conceptual organisation and rearrangement. It is also associated with both improvidence and serialist style, and with globetrotting among scientists. This concentration on facts by rote learning is, however, negatively related to fact retention among scientists. This finding accords with the original work in Gothenburg in which students adopting a surface approach were less able to retain, over a five-week period, factual knowledge from an article they had read than those who had used a deep approach (Marton and Säljö, 1976). This conclusion could also be supported on theoretical grounds from cognitive psychology by arguing that facts are retained more effectively within a framework of meaning which provides opportunities for multiple coding than when stored sequentially with single linkages through rote learning (Cohen, 1983; Entwistle, 1987).

The second factor suggests a substantial overlap between Elaborative Processing, Meaning, Orientation and Positive Attitudes to Studying. Thus the intention to understand seems to involve personalising the information presented and relating it to real-life situations. The third factor is composed mainly of items describing disorganised study methods and social motivation, with elements of negative attitudes.

The fourth factor is less clearly defined. In the total sample it brings together Fact Retention with all three components of Achieving Orientation, but within the faculty groups this tight grouping is dispersed across other factors. Fact Retention contains items which are, in fact, self-ratings of performance ("I do well in exams . . . , I am very good at learning . . . , I have no trouble remembering . . ."). Schmeck (personal communication) now sees this scale as indicating "self-efficacy". This factor thus describes a self-confident, competitive and strategic approach to studying, which is also likely to lead to academic success.

To examine further the constructs underlying the two inventories a factor analysis of the whole 75 items was undertaken (Waterston, 1985). Taking these in conjunction with the earlier item analyses it is possible to suggest groups of items which hold together fairly consistently. The main groupings are organised study methods, analytical thinking, and deep/elaborative processing. But within these groups several conceptually distinct elements could be distinguished which point also to a sequence of attitudes, activities, and cognitive processes (see Table 3).

TABLE 3

INDICATIVE INVENTORY ITEMS

ORGANISED STUDY METHODS

*Purpose/Intention*

— I usually set out to understand thoroughly the meaning of what I am asked to read. (E)
— It's important to me to do really well in the courses here. (E)

*Planning/Strategy*

— I find it impossible to keep to a regular schedule. (—, S)
— When I'm doing a piece of work, I try to bear in mind exactly what that particular tutor or lecturer seems to want. (E)

*Time and Effort*

— I'm rather slow at starting work in the evenings. (—, S)
— I spend less time studying than most of the students in the class. (—, S)

ANALYTICAL THINKING

*Selection/Recognition of Salient Material*

— I have trouble seeing the difference between apparently similar ideas. (—, S)
— I often get criticized for introducing irrelevant ideas. (—, S)

*Organisation of Material*

— Although I generally remember the facts and details, I find it difficult to fit them together into an overall picture. (—, E)
— I have trouble organising the information I have to remember. (—, S)

*Reasoning*

— I have trouble making inferences. (—, S)
— I find it difficult handling questions requiring critical evaluation. (—, S)

DEEP/ELABORATIVE PROCESSING

*Questioning*

— While studying I attempt to find answers to questions I have in mind. (S)
— When I'm reading an article or research report I generally examine the evidence carefully to decide whether the conclusion is justified. (E)

*Relating*

— I try to relate ideas in one topic to those in others wherever possible. (E)
— I learn new concepts or ideas by visualising a situation in which they occur. (S)

*Personalising*

— I like to play around with ideas of my own related to what I am studying. (E)
— After dealing with a new topic, I usually try to summarise it in my own words. (S)

ORGANISED STUDYING

*Remembering/Revising*

— When studying something, I devise a system for recalling it later. (S)
— I use previous exam papers to help me to decide what weight to give to various topics in doing my revision. (E)

Note: E = Entwistle item.
S = Schmeck item.
— = Opposite direction of scoring.

## DISCUSSION

The two inventories were based on quite different theoretical rationales — the one drawn from cognitive psychology and the other from experiential research derived from the qualitative analysis of student self-reports. Both inventories had previously produced four main dimensions. This research suggests a remarkably close coincidence between each set of scales to the point where we seem justified in being able to describe study approaches and processes in terms of four dimensions — deep/elaborative, surface, organised, and strategic/competitive. These dimensions are composed of a spread of items which link, in an intelligible way, intention and motivation with distinctive cognitive processes.

It is important to note also that the factor structure of the learning processes inventory, which was derived initially from students in the USA, has now been confirmed among British students. In other research, the main factors in the Entwistle inventory have been found in Australia (Watkins, 1983), Hungary (Entwistle and Kozéki, 1985), and in Venezuela (Diaz, 1984), although there is a tendency for the achieving and strategic dimensions to merge or be combined with the other two factors in some analyses (see Entwistle, 1988). Using an inventory developed independently of the other two, Biggs (1978, 1987) has reported three main factors — deep, surface and achieving (which combines organisation and competition). There thus seem to be strong reasons for continuing research along these lines, and perhaps creating a single inventory from the strongest items of all three inventories.

A note of caution is perhaps necessary before concluding that this research demonstrates a close correspondence between research on cognitive psychology and student learning. Although the items in the Learning Processes Inventory were derived from research in cognitive psychology, and the main factors extracted parallel the ideas of Craik and Lockhart, the process of extrapolating the research findings to the context of everyday studying is by no means straightforward. The inventory cannot directly operationalise the theoretical constructs, as the theory describes only the mental operations used in tackling narrowly defined experimental tasks. The construction of items represents an imaginative extension of these mental operations into an entirely different context. Yet it must surely be significant that Marton's analysis of students' comments on their approaches to learning led him also to see a "metaphorical resemblance" to Craik and Lockhart's work in characterising the processes he identified as "deep" and "surface" (Marton and Säljö, 1984, page 42).

To what extent does the factor analysis of these self-report inventories begin to provide a plausible description of approaches to studying and processes of learning? In earlier research two main dimensions have repeatedly been found to predict academic achievement — organised study methods and active thinking processes (Weinstein and Underwood, 1985). This research helps to differentiate further the components of organised studying and active thinking in ways which should both facilitate the future measurement of study processes and indicate ways of helping students to develop more effective methods of studying.

There is, however, a danger in focusing too narrowly on the study processes of students as if that studying took place in a vacuum. In fact, the learning environment has profound effects on studying. Approaches to studying are a product of the interaction between the characteristics of individual students and their perceptions of courses, teaching and assessment procedures (Entwistle, 1987). Any attempts to modify students' study strategies are only likely to be effective if the learning environment is also changed in parallel to ensure that the recommended ways of studying are also perceived by students to pay off within the reward structure of the courses.

Acknowledgments.—We are grateful to Dr. R. R. Schmeck for permission to adapt his inventory for use in this study.

Correspondence and requests for reprints should be addressed to Professor N. J. Entwistle, Department of Education, University of Edinburgh, 10 Buccleuch Place, Edinburgh, EH8 9JT.

## REFERENCES

BIGGS, J. B. (1978). Individual and group differences in study processes. *Br. J. educ. Psychol.*, 48, 266-279.

BIGGS, J. B. (1985). The role of metalearning in study processes *Br. J. educ. Psychol.*, 55, 185-212.

BIGGS, J. B. (1987). *Student Approaches to Learning*. Hawthorn, Vict.: Australian Council for Educational Research.

BROWN, W. F., and HOLTZMAN, W. H. (1966). *Manual of the Survey of Study Habits and Attitudes*. New York: Psychological Corporation.

COHEN, G. (1983). *The Psychology of Cognition*. London: Academic Press.

CRAIK, F. M., and LOCKHART, R. S. (1972). Levels of processing: a framework for memory. *J. verb. Learn. verb. Behav.*, 11, 671-684.

DIAZ, D. (1984). The identification of approaches to learning adopted by Venezuelan university students. Unpublished M.Sc. Thesis, University of Wales Institute of Science and Technology, Cardiff.

ENTWISTLE, N. J. (1975). How students learn: information processing, intellectual development and confrontation. *Higher educ. Bull.*, 3, 129-148.

ENTWISTLE, N. J. (1987). A model of the teaching-learning process derived from research on student learning. In RICHARDSON, J. T. E., EYSENCK, M. W., and WARREN-PIPER, D., (Eds.). *Student Learning: Research in Education and Cognitive Psychology*. London: S.R.H.E /Open University Press.

ENTWISTLE, N. J. (1987). *Understanding Classroom Learning*. London: Hodder and Stoughton.

ENTWISTLE, N. J. (1988). Motivational factors in students' approaches to learning. In SCHMECK, R. R. (Ed.), *Learning Styles and Strategies*. New York: Plenum Press.

ENTWISTLE, N. J., and KOZÉKI, B. (1985). Relationships between school motivation, approaches to studying, and attainment among British and Hungarian adolescents. *Br. J educ. Psychol.*, 55, 124-137.

ENTWISTLE, N. J., HANLEY, M., and HOUNSELL, D. J. (1979). Identifying distinctive approaches to studying. *Higher Educ.*, 8, 365-380.

ENTWISTLE, N. J., and RAMSDEN, P. (1983). *Understanding Student Learning*. London: Croom Helm.

ENTWISTLE, N. J., and WILSON, J. D. (1970). Personality, study methods and academic performance. *Univ. Q.*, 24, 147-156.

MARTON, F., HOUNSELL, D. J., and ENTWISTLE, N. J. (1984) (Eds.), *The Experience of Learning*. Edinburgh· Scottish Academic Press.

MARTON, F., and SALJÖ, R. (1976). On qualitative differences in learning. I — Outcome and process. *Br. J. educ. Psychol.*, 46, 4-11.

MARTON, F., and SALJÖ, R. (1984). Approaches to learning. In MARTON, F., HOUNSELL, D. J., and ENTWISTLE, N. J. (Eds.), *The Experience of Learning*. Edinburgh: Scottish Academic Press.

NIE, N. J. *et al.* (1975). *Statistical Package for the Social Sciences*. (2nd edition). New York: McGraw Hill.

PASK, G. (1976). Styles and strategies of learning. *Br. J. educ. Psychol.*, 46, 128-148.

POLLITT, A. B. (1984). Program for Confirmatory Factor Analysis. Edinburgh University: Godfrey Thomson Unit.

RAMSDEN, P. (1981). A study of the relationship between student learning and its academic context. Unpublished Ph.D. thesis, University of Lancaster.

ROHWER, W. D. Jr. (1984). An invitation to a developmental psychology of studying. In MORRISON, F. J., LORD, C. A., and KEATING, D. P. (Eds.), *Advances in Applied Developmental Psychology (Volume 1)*, New York: Academic Press.

SCHMECK, R. R. (1983). Learning styles of college students. In DILLON, R., and SCHMECK, R. R. (Eds.), *Individual Difference in Cognition*. New York: Academic Press.

SCHMECK, R. R., RIBICH, F. D., and RAMANAIAH, N. (1977). Development of a self-report inventory for assessing individual differences in learning processes. *Appl. psychol Measur.*, 413-431.

TAYLOR, E. (1983). Orientations to study: a longitudinal interview investigation of students on two human studies degree courses at Surrey University. Unpublished Ph.D. thesis, University of Surrey.

THOMAS, J. W. (1985). Dimensions of autonomous learning: a framework for investigating attribute-treatment interactions in unsupervised settings. Paper presented at the Annual Meeting of the American Educational Research Association Convention. April, 1985.

WATERSTON, S. (1985). A phenomenal and noumenal approach to student learning compared. Undergraduate dissertation, Department of Psychology, University of Edinburgh.

WATKINS, D. (1983). Assessing tertiary study processes. *Hum. Learn.*, 2, 29-37.

WEINSTEIN, C. E., and UNDERWOOD, V. L. (1985). Learning strategies: the "how" of learning. In SEGAL, J. W., CHIPMAN, S. F., and GLASER, R. (Eds.), *Thinking and Learning Skills (Volume 1)*. Hillsdale, N. J.: Lawrence Erlbaum.

(*Manuscript received* 13*th November*, 1987)

*Br. J. educ. Psychol.*, 58, 266-278, 1988

# DEVELOPING CHILDREN'S THINKING ABOUT THEIR OWN LEARNING

BY INGRID PRAMLING
*(Department of Education and Educational Research, Göteborg University, Sweden)*

SUMMARY. This study focuses on developing children's awareness of their own learning. What children begin to think of as learning, and how they think learning comes about have previously been described in terms of qualitatively different conceptions of learning (Pramling, 1983). In this study, to develop children's awareness means to bring about a qualitative change in how learning is conceptualised. In an experimental-descriptive study three groups of preschool children (in total 56 children) were interviewed, and subsequently observed for two to three weeks while working with the same curriculum topic. After a week they were interviewed a second time, and they were interviewed a third and final time after six months. One of the groups was involved in metacognitive dialogues (Pramling, 1986b) about learning as a natural part of their everyday activities. In this group and one of the others the topic was structured in a similar way. The third teacher did not use either this structure or the metacognitive dialogues. Children's awareness of their own learning increased significantly in the group that worked with dialogues about learning that were related to the content.

## INTRODUCTION

THIS study was concerned with an aspect of metacognition, namely, children's awareness of their own learning. The research approach adopted here differs from the usual approaches to metacognition in that it focused on the child's point of view. The purpose of this exploratory study is to find out if it is possible to develop children's awareness of their own learning. To bring about a change in children's awareness of their own learning means here to change their conception of learning to a more advanced conception. The approach which has been developed in this study is applied in the context of preschool education.

The usual direction within the field of metacognition in the context of education concerns the regulation or control of one's own cognition, that is, metacognitive skills. Children's metacognitive skills are interpreted in relation to information-processing approaches to human thought, in which "the activities of the system are guided by the operations of a central executive, the function of which is to oversee and guide problem-solving" (Reeve and Brown, 1985, p. 347). Such research has focused on finding out what kind of strategies skilled readers use in their learning. It has been found, for instance, that skilled readers use planning, monitoring, and regulating continuously in their reading (Brown, 1978). Many training studies are then aimed at teaching children these strategies which are seen as being applicable in other contexts, for example, in all text reading or problem-solving. The educational implication of this approach to metacognition is to teach children those general strategies. This has been done with greater or lesser success (see e.g., Reeve and Brown, 1985; Paris and Oka, 1986).

In contrast to the skills approach to metacognition, the present study is influenced by phenomenography (Marton, 1981). The point of departure is the relation between the child and the world, which means that metacognition becomes dependent on both content and context as these features appear from the child's point of view. Metacognition is to be seen in three steps. We first focus on *what* the child is thinking about (a content). At the second step we focus on *how* the child is thinking about that content. Finally at the third, we focus on the child's thinking about his own thinking about the content, which is the metacognitive level.

The qualitatively different ways of thinking about learning represent categories on various metacognitive levels. In Pramling (1983) a developmental trend was clearly shown, in the sense that children give expression to some conceptions at an earlier age than others. It seems that children expand their repertoire of ways of thinking, but particular conceptions are not to be found among younger preschool children. From the age of 4 most children have the idea of learning to do, while very few children in preschool and about 28 per cent of the 8-year-olds connect learning with to know about the world. Finally, no preschool children, and only 8 per cent of the 8-year-olds showed an awareness of learning to understand. (In Sweden children begin primary school at 7 years of age, so that in this article the term preschool children refers to those between the ages of 5 and 7.) To develop one's thinking about learning means to change one's way of thinking about learning, from learning to do to learning to know the world or, further, to learning to understand.

From the point of departure of this study, metacognition as dependent on content and context, one has to express some doubts about teaching general strategies to children. First of all, by teaching strategies to children there is always a risk that they develop a way of answering questions which they think the teacher expects of them, and we cannot be sure that they have understood (gained insight into) the phenomenon being studied. Another reason for not teaching children strategies is that understanding is content-dependent. To understand one text does not mean that one understands all texts, since the structure and content differ. To understand a problem in social science does not mean that one can solve a mathematical problem in the same way. As Marton (1986) points out, all attempts to change understanding of something must, of necessity, take their starting-point from how this "something" is constituted. One reason for not teaching children strategies, in other words, is that these strategies do not exist in general terms, but only in relation to a particular content.

The educational implication of the perspective on metacognition adopted in the present study is that the content has to be taught metacognitively. A similar argument has been presented earlier by Biggs (1986) in relation to university students. To work with a content metacognitively means to focus on the child's reflection about that content — to make the child's own thinking an object of his thinking. Children might, for example, be thinking about why it is raining, which is not a metacognitive question. However, if one turns their attention towards how they think about why it rains, it becomes metacognitive. Another example of this approach to teaching is to be found in Martin and Ramsden's (1986) study of how university students changed their conception of learning. The students participated in history classes. As a feature of learning history they had to discuss and reflect on learning in relation to writing essays of their subject of the course over a period of one year.

Doverborg (1987) showed how preschool children perceived what mathematics was all about either as an activity that prepared them for school or as a function of dealing with problems here and now. To see counting as a school activity means that the reason for counting in preschool, from the child's point of view, is because one has to do it in school. On the other hand, children who had a functional conception of counting could talk about a large number of situations in which it is important to be able to count, situations which have meaning for the child. For example, "It is important to be able to count, so that my brother won't cheat me on sweets". Francis (1982) showed how children's approach to learning to read was dependent on their ideas about reading. Children who viewed reading as an activity with personal meaning rapidly learned to read, while those who saw reading only as a school activity without further meaning had more difficulty. Similar findings were made by Dahlgren and Olsson (1985). They showed that almost every child who conceptualised learning to read as a demand from the school did not learn to read during their first year in school, while other children who saw a personal meaning in learning to read did so.

Since it has been shown in the above studies that learning is aided by metacognitive awareness, it would be interesting to find out if metacognitive awareness could be developed in preschool. The purpose of the present study is to explore the feasibility of an attempt to improve metacognitive awareness of learning in preschool children and to assess its effects.

## METHOD

### *Analysis of conceptions of learning*

Children's conceptions of learning were identified by letting them talk about what they had learnt in different contexts, as well as how they had gone about learning that particular skill or phenomenon. In the next step of the analysis the children's points of view were focused on. The statements they made were compared, similarities and differences among these statements were found, and finally the variation in the ways of thinking about the phenomena was described in terms of categories. Children's conceptions of learning were described in two related aspects, namely, what they learn and how this learning comes about. In the most general way, children's conceptions of what they learn were described in terms of "learning to do", "learning to know the world" and "learning to understand". The "how" aspect of learning was described in terms of "learning as doing", "learning by growing older" and finally as "learning by experience" (Pramling, 1983). The categories are described in more concrete terms below.

### *Conceptions of what is learned*

*Learning to do:* The child views learning as the learning of activities, such as manual skills or how to behave. Intellectual skills such as counting, reading, singing or saying words in another language are also included in this category. From the child's point of view these intellectual skills are activities related to articulation, that is, to saying something. All of these activities are connected with the child himself. Examples: "I have learnt to ride my bike", "I have learnt to make drawings of houses", "I have learnt to say 'good-bye' in English", "I have learnt to count to ten", "I have learnt to play shop" or "I have learnt not to shout in the shop".

The child views learning as learning to gain knowledge about something in the world which can be told to someone else. Here the child perceives learning as acquiring information in the mind. The child has gained knowledge of facts or circumstances as a product of someone else's efforts. Example: "I have learnt that the Eiffel Tower is 374 metres high", "I have learnt a lot about elephants that live in the jungle", "I have learnt what it was like in the shop in earlier days" or "I have learnt how cows are stamped after they are slaughtered".

*Learning to understand:* The child views learning as comprehension, that is, the child becomes now aware of its thoughts and can relate a change in its thinking to learning. The child talks about learning as a qualitatively different way of thinking in which it plays an active role itself, that is, the child relates different information or experience and draws conclusions from that. Learning to understand means that learning is seen as a change in way of thinking about reality. The child suddenly sees something in a new light. Examples: "I didn't understand about traffic when I was three, though I knew what cars were", "When I was younger I thought I was the only one with the name Eva, but now I know that many people have the same name", "I thought advertising was just for fun . . . but now I know it is because they want us to buy those things" or "I realised that the money I left to the shop assistant is not only kept at the pay-desk as I thought, but is used for buying new food for the shop".

### *Conceptions of how one learns*

To each of the answers to the question of "what" we learn, there is a corresponding set of qualitatively different answers to the question of "how" we learn. At the first level children do not have any idea that they learn but, instead, perceive it as doing; in other words, learning and doing are the same. At the next level children relate their learning to getting older, that is, when they grow they will be able to do something by virtue of the fact that they become older. Finally, at the third level children realise that learning comes about by experience or a deliberate process. In the present study children only give expression to learning by experience, and for this reason the other categories will not be developed here.

To realise that one learns to *do* something through different kinds of personal experience is a common conception of learning which most children have already developed by the age of 4 (for a detailed description of the origin of the child's view of learning through practice, see Pramling, 1986c). On the other hand, the conception of learning to *know* through one's own experience or activity rarely develops so early. The two categories utilised in the present study in relation to "how" children perceive learning to know will be developed below.

*Learning to know by external influence:* The most common way of perceiving the way in which one acquires knowledge is by "someone telling you" (Pramling, 1983). This means that someone is transferring the knowledge to the child. Examples: "I got to learn about the life of the elephants from my teacher", "I learned about the moon from the astronaut", or "I learned the time by asking my mother repeatedly".

*Learning to know or understand by personal experience:* To understand that one can get to know something from one's own activity means that the child conceptualises situations, settings and activities as sources for acquiring knowledge. The child's own role as a learner is here more active. Examples: "I learned what the shops looked like in olden days by looking at the pictures in the museum", "I learned about how the shop works by playing shop", or "I know what kinds of food to give to my bird by noticing whether he liked what I gave him or not".

### *Interviews*

All the children were interviewed individually one week before they started the theme of "The shop". The interview dealt with questions about learning. (It also dealt with questions about the theme, but that aspect of the study will not be developed in this article.) On the first occasion the question was: "Tell me about something you have learnt in preschool" followed by, "Anything else?" The groups then worked with the theme for two to three weeks, during which they were observed and their comments continuously recorded. On the second occasion the question was: "Now since you have been working on shops, tell me something you have learnt!" followed by, "Anything else?" Finally, the children were interviewed a third time after six months, during which time there was no contact with the preschools. The teachers had not been aware that there would be a follow-up and had carried on with their normal programme. The question then was: "Now since you have been in preschool for a whole year, tell me something you have learnt!"

again followed by, "Anything else?" The interviews were taped and transcribed verbatim.

The interviews were semi-structured, in a way similar to the Piagetian clinical interview, in an attempt to get children to explore their thoughts as far as possible (Doverborg and Pramling, 1985). Observations functioned mainly as a frame of reference for talking with children and understanding their ideas about what they had been involved in. The categories of conceptions described earlier, that is, learning to do, to know and to understand, were utilised in the analysis of the children's ideas about what they learned. The children's responses were categorised according to the highest conception they expressed.

What the children suggested as learning in all the three interviews was then taken as the starting point for asking them about how the learning had come about. Besides the questions concerning "what" and "how" they learned children got questions like: "If you were the one who had to decide what the children will have to learn next in preschool, what would you suggest?", "Imagine you are as old as your teacher, and have to teach children in the next door preschool all you have learnt when working with the shop. How would you go about that?" "Isn't there any other way?" After this, children were asked about why they had done all the different activities, and also if they had learnt anything from going to the shop, making drawings, playing in the shop etc.

The reliability of these categories was tested in the earlier study (Pramling, 1983), and was found to vary between 86 and 100 per cent in the total category system, with a median of 97 per cent (of which the categories presented here are only a part). The inter-judge reliability does not refer to the categorisation as such, that is, to the categories chosen. Reliability is only a measure of the extent to which another person sees the same categories in the interviews as the author does, on the basis of the instructions given. The inter-judge reliability in the present study varied between 93 per cent and 98 per cent for the different categories. Since the difference in the inter-judge reliability is very small the author's judgment prevails.

### *Design and implementation of the study*

Three teachers (A, B, C) worked for two to three weeks with the same content, "The shop". In Figure 1 we can see the design of the study. Two of the teachers (A, B) organised the content in the same way. During one period the children were taught about how to shop (the customer's perspective) and in another period about how the shop functions (the perspective of the shop). Organising the teaching from these two perspectives was based on an assumption that it might be easier for children to understand when there is a clear focus. All the teachers emphasised the customer's perspective as being very important to children. Focusing on the perspective of the shop arose from a study by Jahoda (1979). He showed how children began to understand the two systems of money operating in the shop, i.e., first the relation between the customer and the shop and later the relation between the shop and the supplier. But children up to the age of 11 failed to a great extent to grasp the relation between these two systems of money. The teacher C's performance was not based on any structure and she did not use metacognitive dialogues.

The design makes it possible to compare group A with group B, and group B with group C, to look at the role of metacognitive dialogues, and to compare groups A and B with group C to look at the role of the structure of the teaching.

*Teacher A* carried out "metacognitive dialogues" with the children. This means that she focused the children's attention on different ways of thinking about learning in everyday settings during their work with the shop theme. The children

FIGURE 1

THE DESIGN OF THE STUDY

| Preschool | Content | Structure | Metacognitive dialogues |
|---|---|---|---|
| A | X | X | X |
| B | X | X | |
| C | X | | |

were requested to reflect and ponder about what they were doing and why they were doing certain things which are normally taken for granted, for example, "How come that we went to the shop yesterday?", "Did you find out anything that you didn't know before?", "How did you do about finding it out?" After they had expressed their ideas, she also made them aware of the different ideas they all had about the topic and the ways in which they had found out about it. For example, they began to talk about what it was like in shops in earlier times. The children were given the task of finding out about it by the next day. The next day the children told their friends things about shops long ago, and they also had to say how they found it out. They said they found out by asking parents, grandparents, neighbours, reading in books, looking at a TV programme and so on. The teacher asked them to identify all the different ways in which they had obtained this information about shops. She requested the children to reflect on aspects of learning that are most often taken for granted, such as there being some meaning behind different activities, that one can acquire knowledge from experience, that different activities happening on different occasions are intended to teach the same phenomenon, and why certain activities are repeated.

*Teacher A* taught the content metacognitively by focusing the children's attention on the learning aspect related to the content of the theme. The children were not asked to think about their learning as such, but about their learning in relation to the content. With preschool children this is often a question of concretisation combined with reflection. One example of how this was carried out by Teacher A occurred after the children visited a shopping centre. They had collected and observed a variety of goods, and discussed and looked at their findings. After a while the teacher expressed her satisfaction with all their observations about the shop, advertising and so on. Then they were given a problem to solve: "How would you go about teaching other people all you have learnt about the shop?" The children worked in three groups with the assistance of the teacher. One group wrote (and drew) a large wall-letter, the second made a magazine and the third made a theatre with paper dolls. Afterwards they talked about what led them to present their knowledge in different ways. The children had many suggestions.

*Teacher B* treated the theme with a traditional, school-like way of teaching. She repeated the different parts of the theme over and over, and the children often listened passively. The answers to the questions they received were of a right or wrong type. The teacher continued questioning until the answer she expected came up or until every child had given an answer. There was, for instance, a song about opening a shop and selling something for the cost of 1 krona. Every child had to suggest what to buy. After every answer they sang the song, which means that the song was sung 19 times since there were so many children. Teacher B asked very specific questions such as, "What is the name of the lady whom you pay before you leave the shop?" This was teacher B's normal way of teaching. She taught the content from the customer's and the shop's point of view, but she did it by labelling facts repeatedly or asking children to express their knowledge.

*Teacher C* was only asked to work on "The shop" the way she wanted to. Her teaching covered the same content as the other two teachers' but it was not organised into the customer's or the shop's perspective. This teacher's way of working with children was a traditional preschool method developed from Fröebel's ideas of young children's learning. According to this approach, learning takes place when teachers provide materials for play or show children how to construct something which everybody then has to practise. In Teacher C's classroom, learning took place via play and exercises in terms of doing things in clay, cutting things out of paper, making constructions and so on. The children were manually very active. They played and constructed many things but the teacher did not encourage the children to reflect on what they were doing or why.

### *The children*

As we can see in Table 1, three groups of preschool children and their teachers participated. In total there were 56 children between 5 and 7 years of age. The groups had approximately the same number of children, with boys predominating in two of them. Group C had the youngest age group, with no child having been in preschool for more than one year, while almost half of the children in group B were in their second year. In every group there was a nursery nurse in addition to the preschool teacher. All three preschools were located in middle-class areas in a large city, which means that the children's backgrounds were comparable.

TABLE 1

NUMBER AND DISTRIBUTION OF INTERVIEWED CHILDREN

| | Group | | |
|---|---|---|---|
| | A | B | C |
| Number of children | 20 | 19 | 17 |
| Girls | 9 | 7 | 9 |
| Boys | 11 | 12 | 8 |
| 5 years old | 4 | 6 | 10 |
| 6 years old | 16 | 13 | 7 |
| Mean age | 5:11 | 5:11 | 5:8 |
| Variation | 5:0-6·8 | 4:9-6:6 | 5:1-6:10 |
| In preschool one year before | 3 | 9 | 0 |

These figures are from the first interview.

The groups means showed no significant difference in age. However, in group B half of the children had spent a longer time in preschool.

## RESULTS

The results consist of a categorisation of answers given by children in the three different groups on the three different interview occasions into the categories described earlier.

### *The children's conceptions of what they learn*

As we can see in Table 2, at the first interview there is little difference among children's conceptions of what they have learnt. Most children expressed learning to do, and there was no significant difference between the groups ($\chi^2 = 4{\cdot}3$, df = 4, $P > 0{\cdot}05$). However, at the second interview there is a significant difference ($\chi^2 = 18{\cdot}56$, df = 4, $P < 0{\cdot}01$). The table shows that ¾ of the children in the "experimental group" (A) now answered differently after studying the theme.

TABLE 2

CHILDREN'S CONCEPTIONS OF WHAT THEY HAVE LEARNT BEFORE AND AFTER COMPLETING THE SHOP THEME

| | Group | | | | | |
|---|---|---|---|---|---|---|
| | A | | B | | C | |
| Occasion | 1 | 2 | 1 | 2 | 1 | 2 |
| Do not know | 3 | 3 | 1 | 5 | 2 | 4 |
| Learning to do | 16 | 1 | 17 | 11 | 13 | 8 |
| Learning to know | 0 | 14 | 1 | 3 | 2 | 4 |
| Learning to understand | 1 | 1 | 0 | 0 | 0 | 0 |
| Number of children | 20 | 19 | 19 | 19 | 17 | 16 |

### *What children say they would like to learn*

Another question concerned with exploring what the child perceived as learning was: "If you were the one who had to decide what the children will have to learn next in preschool, what would you suggest?" The answers which have been categorised within learning to do focus on the same kind of content as in the study already described, that is, to draw, to walk on stilts, to use socks indoors and so on. Examples of answers which have been categorised within learning as knowing are "I want to know how the first ape men grew up", "I want the teacher to tell us who invented the telephone" and "I want to learn how to tell the time". No child expressed anything that was categorised within learning as understanding. The reason for this could be that this was the last question in the interview and was not followed up as the other questions were. Further, this question was only asked in the second interview. However, as we can see in Table 3, more children from the group A suggested what to learn in a way that was categorised within learning as knowing. The chi square test shows a significant difference between the groups ($\chi^2 = 6{\cdot}61$, df = 2, $P < 0{\cdot}05$), with the largest deviation from the expected value in group A. Learning as knowing is a conception which rarely occurs spontaneously among preschool children. It was earlier shown in Pramling (1983) that this is a conception which is only held by about 15 per cent of children in the age range 5 to 6 years.

TABLE 3

WHAT CHILDREN SUGGEST AS A CONTENT OF FURTHER LEARNING IN PRESCHOOL

| | Group | | |
|---|---|---|---|
| | A | B | C |
| Do not know | 4 | 7 | 4 |
| Learning to do | 7 | 11 | 9 |
| Learning to know | 8 | 1 | 3 |
| Number of children | 19 | 19 | 16 |

What we have seen here is that children in group A seem to have become aware of other aspects of learning, where they reflected about what they had learnt. This increase in awareness was not evident in the other two groups.

*Children's conceptions of how learning comes about*

Let us now look at how children express their conception of how learning comes about. One question was: "Imagine you are as old as your teacher, and have to teach children in another preschool all that you have learnt when working with the shop. How would you go about that?" A similar question was asked in an earlier study where children had worked with another curriculum topic (Pramling, 1986a). In that study all the children who had some idea of how to teach someone else said that they should tell the other children about it. As we can see in Table 4, the results of the present study show that children expressed two conceptions, namely, either they should teach someone else by telling about the shop, or they should teach someone else by giving them experience. Giving children experience could be taking children to the shop, looking in magazines for food, working with art and crafts, and so on. Earlier research showed that the latter conception, that is, the idea that one can learn from experience (in the sense of knowing), is not evident to most preschool children. It is notable that so many children in group A expressed it.

TABLE 4

CHILDREN'S CONCEPTIONS OF HOW TO TEACH SOMEONE ELSE (SECOND INTERVIEW)

| | Group | | |
|---|---|---|---|
| | A | B | C |
| Do not know | 4 | 11 | 7 |
| External influence | 4 | 7 | 7 |
| Person experience | 11 | 1 | 2 |
| Number of children | 19 | 19 | 16 |

In Table 4 we can see that there is also a significant difference among the groups with respect to how children conceptualise the way learning comes about ($\chi^2 = 10{\cdot}20$, df = 2, $P < 0{\cdot}01$). More than half of the children in group A suggested that one can teach by letting the children have some experience in terms of activities, while only a few other groups came up with this idea.

The children were also asked about the aim of all the different activities which they had been involved in during the shop theme, such as class discussion, visits to the shop, cutting out pictures from magazines, building a shop and "playing shop". They were, in other words, asked about "Why did you . . ." as well as "Did you learn anything from . . .?". It should be pointed out that all these activities are, from the teacher's point of view, the accepted methods for teaching children about the shop, since she has learnt in her training course that children learn from concrete experience. The teacher takes for granted that all these activities are bricks in building up knowledge about the shop. It appeared from the interviews that children conceptualised the aim of these activities in two qualitatively different ways; either they saw the activity as a goal, or they saw it as a means.

Children who conceptualised the activity as a goal recognised what they had been doing, the activity as such, but not any further aim. It seemed that the children took for granted that they did certain things that the teacher told them to or because they liked it, for example, "Because it is fun to play in a shop", "Because the teacher asked us to cut". Other children, on the other hand, conceptualised the activity as having a purpose, that is, that the activity was a means of teaching them something. These children were able to see something behind the situation — the teacher's intentions. For example, "We went to the shop to buy things and to learn where you can find different things". The categorisation in Table 5 has been made from a hierarchical point of view. This implies that conceptualising the activity as a means is a more advance conception, as shown in earlier studies (Pramling, 1983).

TABLE 5

CHILDREN'S CONCEPTIONS OF THE DIFFERENT ACTIVITIES THEY HAVE BEEN INVOLVED IN (SECOND INTERVIEW)

| | Group | | |
|---|---|---|---|
| | A | B | C |
| The activity as a goal | 5 | 11 | 8 |
| The activity as a means | 14 | 8 | 8 |
| Number of children | 19 | 19 | 16 |

The chi square test shows no significant difference among the groups ($\chi^2 = 4{\cdot}12$, df = 2, $P > 0{\cdot}05$). The reason that no systematic differences were found may be that the children were asked about many different activities. If they saw the activity as a means to something else in relation to *one* of these activities, their response was categorised as a conception of the activity as a means.

### *The third interview*

### *The "what" aspect of learning*

As we can see in Table 6, the differences between the group A and the other two groups were even greater after 6 months. With respect to how children conceptualise learning in preschool there was only one child in group A who did not go beyond the conception of "learning to do", while none of the children in group B came up with anything beyond this. In group C three children expressed learning to know, while ¾ of the children in group A did so. Finally, we can also see that two children now expressed learning to understand. The chi square test shows a significant difference between the groups ($\chi^2 = 31{\cdot}34$, df = 4, $P < 0{\cdot}01$).

TABLE 6

CHILDREN'S CONCEPTIONS OF WHAT THEY HAVE LEARNT DURING THE YEAR IN PRESCHOOL (THIRD INTERVIEW)

| | Group | | |
|---|---|---|---|
| | A | B | C |
| Do not know | 0 | 2 | 4 |
| Learning to do | 1 | 15 | 9 |
| Learning to know | 14 | 0 | 3 |
| Learning to understand | 2 | 0 | 0 |
| Number of children | 17 | 17 | 16 |

To write in a book (that is, fill in simple exercises) is something which many children in groups B and C expressed as an example of what they had learnt during the year. This means that their conception was that they had learnt to do something, "To keep the pencil on a line when drawing" or "To draw nicely", for example. Many of these children seemed to connect their writing with a school activity; it was something one had to do in school. The writing or drawing was the goal. When children in group A talked about their books (which they had made themselves) they expressed having learnt about something else by working with their books. The writing as such was not the goal, but a means of learning a content, "By thinking about the bee and drawing it in my book I found out about the pollen sacs on the bee's legs", for example. The children's ideas of why they worked with their books, in other words, are reminiscent of how they conceptualised the activities described earlier in connection with the theme of the shop. Some children saw the activity as such, while others perceived that the activity was intended to teach them about something.

### *The "how" aspect of learning*

A follow-up question to what they had learnt in preschool was: "How did you go about learning that (which refers to what the child is talking about)?" The children who talked about learning to do something all referred to some kind of experience, either in the form of the teacher's instruction (showing or telling) or by doing it themselves. The fact that one learns to *do* something through experience was thus obvious to these children. All of these advanced at least to the stage of stating that they had learned to know through personal experience. It was pointed out earlier that it is considerably more difficult to understand that one can also acquire knowledge, learn to *know*, through experience. Table 7 relates only to the children who answered the "what" question with examples of learning to know or to understand. That is why no children from group B and only three children from group C appear. It is also the reason for not performing a chi square test.

TABLE 7

CHILDREN'S CONCEPTIONS OF HOW TO LEARN TO KNOW OR TO LEARN TO UNDERSTAND OCCURS (THIRD INTERVIEW)

| | Group | | |
|---|---|---|---|
| | A | B | C |
| External influence | 0 | 0 | 3 |
| Person experience | 11 | 0 | 0 |
| Reflection | 5 | 0 | 0 |
| Number of children | 16 | 0 | 3 |

As we can see in Table 7, there are marked diferences between the groups. None of the children in group A stopped at the category learn to know by being told something, or by external influence. In this group we can also see a new category labelled "reflection". Five of these children now expressed that they had learnt to know by reflecting which can be considered a very advanced conception for children of this age. One example is, "I suddenly realised that people take away the bee's food, when I thought about what the teacher told us".

## DISCUSSION

As may be seen, the investigation has the form of an incomplete cross-experimental design in which two independent variables (the structure of teaching the content and metacognitive dialogues) are introduced. It should be emphasised, however, that given the nature of the phenomena many variables are uncontrolled and must therefore be taken into consideration in discussing the significance of the findings. Indeed, the study should be seen as preliminary to developing a larger and more controlled study in the future.

A comparison of the children's answers to what they learned and how they thought it came about from the first interview showed no differences between the groups. During the intervention, the teaching in group A was built on metacognitive dialogues and clear structure, while the teaching in group B was built on a clear structure without such dialogues. The teaching in group C was not built on either metacognitive dialogues and/or a clear structure. In comparing groups B and C it appears that the structure did not help the children to learn about their own learning. Of course, it is impossible to prove that the development took place as a consequence of metacognitive dialogues. Establishing this would require a fourth group which only worked with these dialogues. So, one cannot exclude the possibility that the increase in awareness in group A was due to a combination of the structure and the metacognitive dialogues. It is important to stress that, in spite of the experimental design, it is of great importance that the study should be seen as a feasibility study, that is, as a study concerning the development that is possible as a result of educational experience.

What we have seen in this study is that a large proportion of the children who participated in the experimental group developed an awareness of their own learning. This development did not occur as a consequence of training any general strategies, but as a consequence of changing perspective in relation to particular activities. When children think, their thinking is always focused on something in the world. By getting children to change their perspective, they reflect on how they think about the phenomena in the world. It seems that the differences between group A and the other two groups in terms of awareness of their own learning increased further during a six-month period.

The most probable reason for the expansion of children's thinking in group A is that this teacher continuously focused on the children's ideas of learning throughout the year. In an interview, she said that she continued to ask the children other kinds of questions and to think about why they were doing different activities. One example she told us about happened one day when the group was passing by a garden and discovered an unusual flower. The teacher asked if anyone had any idea about its name. Since no one knew, she asked how they could find it out. The children had a lot of suggestions, such as, walking into the house and asking, asking their mothers, finding it in a book, giving it a name themselves, and calling the question programme on the radio. The teacher also made them aware of the variation in the different ideas they had about how to find it out. Finally the teacher decided that they should find the name of it for the next day as homework. The following day eight children came with the name of the flower, and they all had to tell the others how they had learnt it.

Different perspectives on metacognition give rise to different research approaches and educational implications. In an educational context one can always choose different perspectives. What has been argued for here is to teach the content metacognitively, which is different from teaching strategies or from teaching facts to children. The results are sufficiently interesting to motivate further and better controlled studies.

ACKNOWLEDGMENTS.—The research reported here was financially supported by a grant from the Swedish Council for Research in the Humanities and Social Sciences.

I express my deep thanks to Professor Melvin Feffer, Rutgers University, and Professor Ference Marton, University of Göteborg, for their valuable contributions to the writing of this article. Gloria Dall'Alba has checked my English which I also appreciate very much.

Correspondence and requests for reprints should be addressed to Dr. Ingrid Pramling, Department of Education and Educational Research, Goteborg University, Box 1010, 431 26 Mölndal, Sweden.

## REFERENCES

BIGGS, J. (1986). Enhancing learning skills: the role of metacognition. In BOWDEN, J. A. (Ed.), *Student Learning: Research into Practice.* Centre for the Study of Higher Education, University of Melbourne.

BROWN, A. (1978). Knowing when, where and how to remember: a problem of metacognition. In GLASER, R. (Ed.), *Advances in Instructional Psychology, Vol. 1.* Hillsdale, N.J.: Lawrence Erlbaum.

DAHLGREN, G., and OLSSON, L-E. (1985). *Läsning i Barnperspektiv.* (Reading from the perspective of the child.) Göteborg: Acta Universitatis Gothoburgensis.

DOVERBORG, E. (1987). *Forskolebarn och Matematik?* (Preschool children and mathematics.) Publikation från Institutionen for pedagogik, Göteborgs Universitet. Nr. 5.

DOVERBORG, E., and PRAMLING, I. (1985). *Att förstå barns tankar.* (Understanding children's thinking.) Metodik for barnintervjuer. Stockholm: Liber.

FRANCIS, H. (1982). *Learning to Read.* London: George Allen and Unwin.

JAHODA, G. (1979). The construction of economic reality by some Glaswegian Children. *Eur. J. soc. Psychol.*, 9, 115-127.

MARTIN, E., and RAMSDEN, P. (1986). Do learning skills courses improve student learning? In BOWDEN, J. A. (Ed.), *Student Learning: Research into Practice.* Centre for the Study of Higher Education, University of Melbourne.

MARTON, F. (1981). Phenomenography — describing conceptions of the world around us. *Instruct. Sci.*, 10, 177-200.

MARTON, F. (1986). Some reflections on the improvement of learning. In BOWDEN, J. A. (Ed.), *Student Learning: Research into Practice.* Centre for the Study of Higher Education, University of Melbourne.

PARIS, S., and OKA, E. (1986). Children's reading strategies, metacognition and motivation. *Dev. Rev.*, 6, 25-56.

PRAMLING, I. (1983). *The Child's Conception of Learning.* Goteborg: Acta Universitatis Gothoburgensis.

PRAMLING, I. (1986a). *Barn och Inlarning.* (Children and learning.) Lund: Studentlitteratur.

PRAMLING, I. (1986b). *Meta-inlarning i förskolan.* (Meta-learning in preschool.) Forskolebarns uppfattningar av inlarning och innehållet i temat "Affären". Publikation från Institutionen for pedagogik, Goteborgs Universitet. Nr. 15.

PRAMLING, I. (1986c). The origin of the child's conception of learning through practice. *Eur. J. Psychol. Educ.*, 1 (3), 31-46.

REEVE, R., and BROWN, A. (1985). Metacognition reconsidered: implications for intervention research. *J. abnorm. Child Psychol.*, 13, 343-356.

(*Manuscript received* 19*th August*, 1987)

*Br. J. educ. Psychol.*, 58, 279-286, 1988

# INTEGRATION OF DOWN'S SYNDROME CHILDREN IN THE PRIMARY SCHOOL: A LONGITUDINAL STUDY OF COGNITIVE DEVELOPMENT AND ACADEMIC ATTAINMENTS

By WENDY CASEY[1], DAVID JONES[2], BARBARA KUGLER[3]
AND BRIAN WATKINS[4]
[1] *(Department of Clinical Psychology, Harperbury Hospital, Radlett)*
[2] *(Department of Psychology, Birkbeck College, University of London)*
[3] *(Harper House Children's Service, Harperbury Hospital, Radlett)*
[4] *(Haringey School Psychological Service, Tetherdown, London)*

SUMMARY. Thirty-six Down's Syndrome children, half of whom attended mainstream primary or nursery schools and half schools for children with moderate learning difficulties, were assessed over a two-year period. Baseline and six-monthly measurements were taken of their progress in expressive language, comprehension, numeracy, verbal fluency, drawing ability and reading. Mental age was estimated at baseline and at the end of the two years. The mainstream children made significantly greater progress in numeracy, comprehension and mental age and compared favourably on all other measures. The results are discussed in terms of their implications for integrated education.

IN the period since the Warnock Committee Report (1978) there has been an increased interest in the problems involved in the integration of children with special educational needs into mainstream schools. A number of studies sponsored by the Schools Council and the National Foundation for Educational Research have focused on the range and nature of adjustments in teaching practice, curriculum and the physical environment as represented in the school buildings to enable mainstream schools to provide for children with special needs (Hegarty and Pocklington, 1981; Hegarty, 1982; Hegarty *et al.*, 1982; and Hodgson *et al.*, 1984). A willingness to be flexible and to experiment with placements seems to have been more noticeable in cases of children of near average ability with sensory or physical handicaps than for those with moderate or severe learning difficulties. In addition, a great deal of effort has gone into the provision of extra support facilities to enable mainstream schools to contain children with emotional difficulties and conduct disorders.

A similar trend towards integration and containment in mainstream schools has been under way in the United States for some years although the motivation to reduce segregation has not always been for the same reasons as in this country. Much of the pressure for change in the United States followed the more general demands for equality of treatment in education for minority groups (Zigler and Hall, 1986; Yoshida, 1986). Part of the concern in the United Kingdom has been over the seemingly ever increasing numbers of children being placed in special schools although as Brennan (1982) points out there are many who still support concentration of specialised resources as an effective way of educating handicapped children. Zigler and Hall (1986) caution against "normalisation" becoming a treatment fad without a scientific basis. Stobart (1986) emphasises the point that the move towards integration has not been based on changes in educational or psychological theories of learning. Physical contact between children with special needs and the supposedly normal will not on its own be sufficient to produce or explain changes in academic attainment and social interaction. There are almost as many degrees and forms of integration as there are categories of special needs. It is much too early to make value judgments about success and failure of integration. Perhaps the most important single factor has been the degree of willingness of the mainstream school to accept or retain the child with special needs. Also important is an acceptance that in special cases the goals of education may have to be modified to minimise the effects of a disability (Wall, 1979).

In the case of children with Down's Syndrome much of the initial pressure to make mainstream education available has come from parents. Some Education Authorities have shown an early readiness to respond to demands at least at the level of nursery and infant school placements. Case reports indicate that some of these placements are successful, e.g., Lorenz (1983). More general evidence on the results of mainstreaming children of below average ability is reviewed by Corman and Gottlieb (1978). There remains a need for more longitudinal data.

The present study is an attempt to monitor and evaluate the cognitive development and academic attainments of Down's Syndrome children over a two-year period. It was thought that a one-year study would be too short to give reliable data especially for the younger children, some of whom were still attending nursery schools at the start of the investigation. Half of the children attended mainstream schools and the others were placed in special schools. This level of comparison does not allow a definitive evaluation of the different education systems involved; nevertheless it is possible to consider whether at least some of the Down's Syndrome children placed in mainstream schools make as much progress as those receiving special education.

## METHOD

### *Sample*

Thirty-six Down's Syndrome children, half of whom attended mainstream schools and half attended schools for children with moderate learning difficulties (MLD), were identified by approaches to Local Education Authorities (LEAs) in Greater London and the Home Counties. Initially the age range 4 to 7 years was specified but as this produced too few children the limits were extended to include children between 3 years 8 months and 10 years of age. Initially there were 11 girls and 10 boys in the mainstream sample and 9 girls and 10 boys in the MLD sample. In the mainstream group one parent refused permission for her daughter to take part and during the two-year study one girl and one boy changed to MLD education after the first year and had to be excluded from the sample. In the MLD group one boy was excluded during the two-year study because a previously undetected hearing loss was corrected and would have confounded the changes in his test performances. The final sample thus produced equal numbers of boys and girls in both mainstream and MLD schools.

In all, 12 MLD and 17 mainstream schools across 11 LEAs took part in the study. In four LEAs the general policy was to place MLD Down's children in MLD schools and only one child in these LEAs was attending a mainstream school; this was at the parents' specific request. In six LEAs the general policy was to place MLD Down's children in mainstream schools and all children from these LEAs in our sample were thus placed. In the remaining LEA parents had the option of mainstream/MLD placement for their Down's children. Three children from this LEA were included in the present sample. By parental preference, two children were in mainstream education and one was in MLD education.

Four children in the total sample were of non-white or mixed parentage, but English was the first language in all homes. Social variables such as ethnic origin, family size, parental education or social class were not taken into account in the selection process, but it was noted that the social class distributions for the two samples were similar (Office of Population Censuses and Surveys, 1970). Educational variables such as class size, curriculum and presence or absence of additional teaching support were not controlled. Class size in the mainstream schools ranged from eight to 34 and in the MLD schools from eight to 14. Descriptive data on the children are shown in Table 1.

TABLE 1

MEANS AND RANGES (IN MONTHS) FOR CHRONOLOGICAL AND MENTAL AGES (C.A.; M.A.) ACCORDING TO SCHOOL (MAINSTREAM AND MLD) AND SEX

| | | Mainstream | | MLD | |
|---|---|---|---|---|---|
| | | Girls (N = 9) | Boys (N = 9) | Girls (N = 9) | Boys (N = 9) |
| CA | Mean | 72·44 | 62·89 | 82·67 | 80·67 |
| | Range | 44-119 | 46-109 | 61-95 | 59-99 |
| MA | Mean | 48·89 | 40·89 | 44·11 | 42·33 |
| | Range | 29-80 | 27-74 | 36-53 | 29-51 |

### *Procedure*

Testing was carried out by the first author at either the child's home or school. Each child was studied over a period of two years and was tested on five occasions: baseline, 6, 12, 18 and 24 months. Six-monthly test intervals were chosen to give detailed monitoring of progress without making undue demands on the children's time.

### *Measures*

The children's progress was evaluated using standardised intelligence, language, ability and reading tests of the kind commonly employed to monitor these factors in the normal population and in other studies of Down's Syndrome children (Lorenz *et al.*, 1985). Mental Age (MA) and Intelligence Quotient (IQ) were measured at baseline and at the final 24 months assessment using the Stanford-Binet Intelligence Scale (Form L-M, 1960) with the 1977 Mental Age score corrections (see Sattler, 1982). The Binet Scale was chosen because it provides a wide enough range to measure the ability of even the youngest children in the study.

On all five test occasions the following measures were obtained from each child:

*Language comprehension* (CL) and *Expressive language* (EL). These were measured using the Reynell Developmental Language Scales (RDLS) which provide a relatively quick means of assessing language development.

*Numeracy* (NUM), *Verbal Fluency* (VF) and *Drawing Ability* (DR). These functions were evaluated using selected scales from the McCarthy Scales of Children's Abilities (MSCA). The MSCA was chosen for similar reasons to the Binet, namely that the scales cover a wide range of ability and allow progress to be measured in months. NUM was selected because of its specific relevance as a measure of scholastic attainment. VF is a measure of the ability to classify and think categorically. DR was selected as a measure of non-verbal motor skill, since modelling, drawing and cutting have been noted as important components of the early, active learning process involving symbolisation of objects and events (Piaget, 1963; Cratty, 1972). The MSCA scales selected for this study were Number Questions, Counting and Sorting, Verbal Fluency, Draw-a-Design and Draw-a-Child. Where more than one scale was used to generate a test age, the mean test age was taken. Test ages were extrapolated from Table 17 of the MSCA Manual (pp. 204-205).

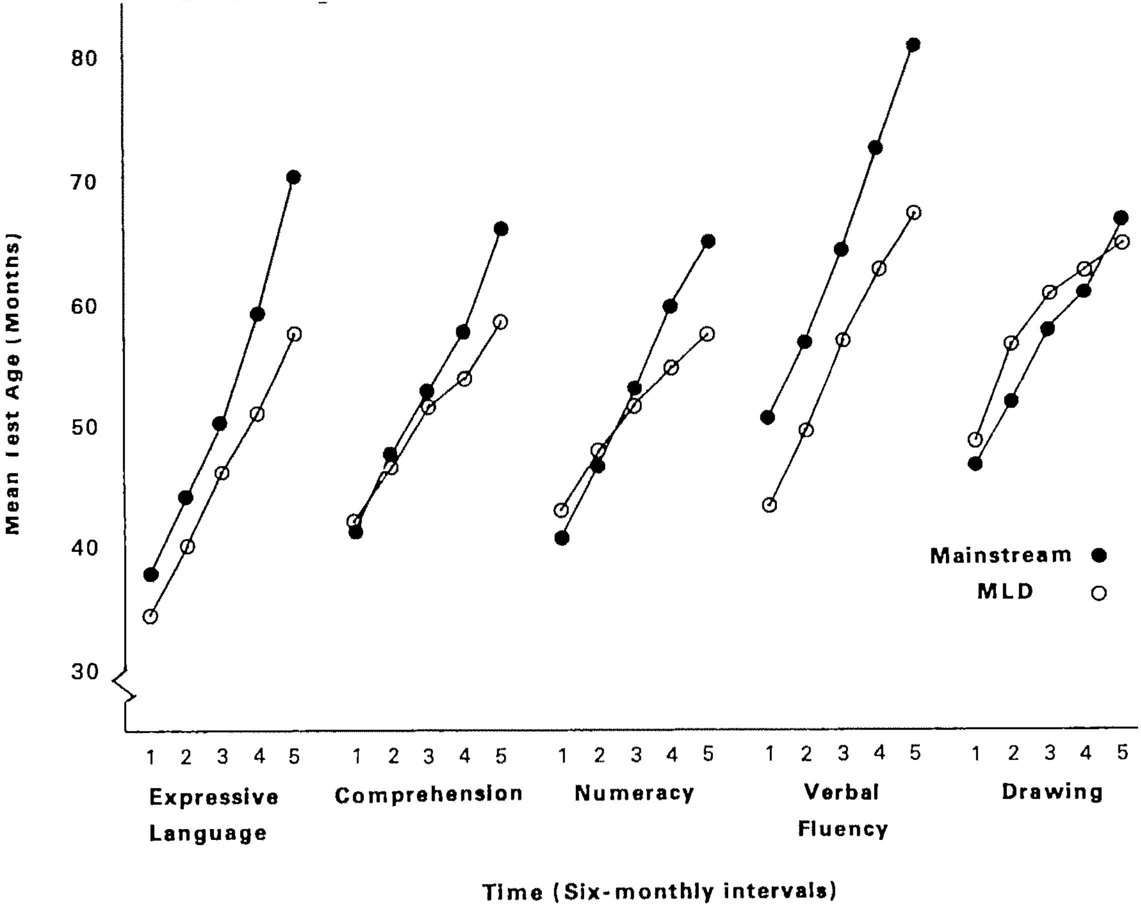
*Reading attainment.* The Neale Analysis of Reading Ability was used to measure progress in word recognition and reading comprehension. The format of the test is similar to that of reading material used in classrooms with children of this age.

## RESULTS

Mean test ages for all five RDLS and MSCA measures on each of the five test occasions are shown by type of school in Figure 1. Each child's six-monthly successive results on the RDLS and MSCA measures showed substantial agreement

FIGURE 1

Mean Test Ages for Each RDLS and MSCA Measure on Each of the Five Test Occasions



with fitted straight lines using the standard least squares procedure and, as a consequence, it was felt that further analysis could be restricted to the initial (Occasion 1) and final (Occasion 5) scores on all these measures.

The effects of schooling and sex of the child on initial EL, CL, NUM, VF and DR scores were assessed by two-way analyses of variance. In all these analyses, the only significant effect was that the EL scores for girls were better than those for boys at the beginning of the study (F (1,32) = 5·46, P < 0·05).

Similar school × sex analyses were made of initial chronological and mental ages (see Table 1). The MA analysis showed no significant differences between mainstream and MLD groups, between girls and boys, nor any significant interactions. With the analysis of the CAs the only significant effect indicated a higher CA for MLD than for mainstream children (F (1,32) = 5·026, P < 0·05). On the basis of clinical experience it was felt that this CA difference had not arisen as a result of any administrative selection procedure but by chance as children of all ages attended both mainstream and MLD schools. In fact, as Table 1 indicates, while the mean CA was higher in the MLD group than in the mainstream group, the age ranges of both boys and girls in the MLD schools lay within the ranges in the mainstream schools.

Overall, initial values on all measures (the five RDLS and MSCA measures and MA and CA) varied considerably within each school group and it was felt that a clearer test of differences would be obtained if this variance was taken into account. Thus, analysis of covariance was used to assess differences between mainstream and MLD children's final (Occasion 5) RDLS and MSCA scores, adjusted for possible differences in CA, MA and the relevant initial (Occasion 1) RDLS and MSCA scores.

The final scores, adjusted by analysis of covariance, showed no significant effects either for sex or for the interaction of sex and school on any of the five RDLS and MSCA measures. There were, however, significant effects indicating that children in mainstream schools obtained higher numeracy scores on Occasion 5 than children in MLD schools (F (1,31) = 7·558, P < 0·025) and, similarly, that comprehension scores on Occasion 5 were higher for mainstream than for MLD children (F (1,31) = 4·398, P < 0·05). Table 2 shows RDLS and MSCA Occasion 1 scores and Occasion 5 original and adjusted scores.

TABLE 2

MEAN INITIAL, FINAL AND ADJUSTED FINAL TEST AGES (IN MONTHS) FOR EACH OF THE FIVE RDLS AND MSCA MEASURES (EL, CL, NUM, VF AND DR), TABLED ACCORDING TO SCHOOL (MAINSTREAM AND MLD)

| | Occasion 1 | | Occasion 5 | | | |
|---|---|---|---|---|---|---|
| | Initial Score | | Final Score | | Adjusted Score | |
| | Mainstream | MLD | Mainstream | MLD | Mainstream | MLD |
| EL | 37·72 | 34·44 | 70·03 | 57·33 | 66·55 | 59·02 |
| CL | 43·11 | 43·67 | 65·33 | 58·30 | 66·12 | 57·80 |
| NUM | 40·44 | 42·83 | 64·72 | 57·06 | 65·28 | 56·50 |
| VF | 50·28 | 42·78 | 80·56 | 66·94 | 72·97 | 74·53 |
| DR | 46·72 | 48·39 | 66·67 | 64·72 | 65·69 | 65·70 |

In addition, while analysis of variance of MA scores obtained at the beginning of the study showed no differences between mainstream and MLD school groups, a two-way analysis of variance on gains in MA over the two years of the study showed a significantly greater gain for mainstream than for MLD children, the former increasing by an average of 19·00 months and the latter by an average of 14·17 months (F (1,32) = 7·966, P < 0·01). No other main or interaction effects were significant. The initial (Occasion 1) and final (Occasion 5) MAs of the mainstream and MLD children are shown in Table 3.

TABLE 3

MEANS AND STANDARD DEVIATIONS (IN PARENTHESES) OF INITIAL AND FINAL MENTAL AGES (MA I AND MA II) OF THE MAINSTREAM AND MLD CHILDREN

| | Mainstream | MLD |
|---|---|---|
| MA I | 44·89 (15·67) | 43·22 (6·36) |
| MA II | 63·89 (13·86) | 57·39 (6·84) |

Within the subject groups there were insufficient numbers of children able to read to permit similar statistical analyses to those used with the other measures. However, the percentages of children reading well enough to obtain scores at the beginning of the study, after one year and after two years (Occasions 1, 3 and 5) are reported in Table 4 together with the percentages of each subject group reading with comprehension.

TABLE 4

PERCENTAGE OF MAINSTREAM AND MLD GIRLS AND BOYS OBTAINING READING SCORES AT BASELINE, 1 AND 2 YEARS

| | Reading Accuracy | | | | Reading Comprehension | | | |
|---|---|---|---|---|---|---|---|---|
| | Mainstream | | MLD | | Mainstream | | MLD | |
| | Girls (N = 9) | Boys (N = 9) | Girls (N = 9) | Boys (N = 9) | Girls (N = 9) | Boys (N = 9) | Girls (N = 9) | Boys (N = 9) |
| Baseline | 44·44 | 22·22 | 22·22 | 11·11 | 33·33 | 11·11 | 22·22 | 0 |
| 1 year | 66·67 | 44·44 | 77·78 | 33·33 | 55·55 | 22·22 | 33·33 | 0 |
| 2 years | 88·89 | 66·67 | 88·89 | 33·33 | 88·89 | 66·67 | 44·44 | 33·33 |

## DISCUSSION

It is apparent that there is an overall trend within the results which would suggest that Down's Syndrome children in mainstream schools do at least as well as, and frequently better than, their peers in MLD schools. It can be seen from Figure 1 that in general the children in mainstream education perform better on all the variables measured over the two-year period.

The progress made by the mainstream children is not attributable to differences between the groups at or before baseline as only one difference at this stage was significant; all the girls were better on expressive language than the boys. This initial sex difference had disappeared two years later. Factors such as later maturation of language skills in boys (Rebelsky *et al.*, 1967; Cunningham *et al.*, 1985) may have been involved. The main differences then between the mainstream and MLD children in terms of the effects of schooling are in higher numeracy and comprehension scores and a significantly greater gain in MA in the former group at the end of the two-year period.

*Numeracy and Comprehension.* The observation that mainstream schools seem to facilitate the development of numeracy skills to a greater extent than special schools has already been shown in other studies with non-Down's Syndrome MLD children in the USA (Bradfield *et al.*, 1973; Corman and Gottlieb, 1978). It would suggest that gains in the ability to carry out simple numerical operations (e.g., counting with and without aids, sorting by set and the like) are more likely to be achieved in an ordinary school class than in the more specialised setting of the MLD school. This finding is particularly important in view of the observation made by Gibson (1978) in his review of the ability profiles of Down's Syndrome children that number sense is usually described as poor when compared with other academic skills.

The small, but significantly greater, improvement in the comprehension scores of the children in mainstream schools relative to those in MLD schools may reflect their wider exposure to verbal interaction with a linguistically more able peer group.

*MA Scores.* The Stanford-Binet MA is a composite score which represents an overall level of achievement encompassing expressive language, comprehension, verbal reasoning, perceptual matching skills and other areas of cognitive ability. The MA provides a general index of ability without reflecting the child's strengths and weaknesses in cognitive functioning since all ipsative information is lost. It is important to note that the mainstream Down's Syndrome children make significantly greater gains than do their peers in MLD schools (19 months as opposed to 14 months). It is perhaps surprising that schools with curricula specifically designed for children with special educational needs do not necessarily facilitate better improvements in general cognitive ability than do the ordinary primary schools.

*Reading Skills.* There is a differential frequency of reading skills in the two sex groups, as well as for the two educational environments, although the relation between type of schooling and reading ability is not as clear-cut as that found by Carroll (1967) or Bradfield *et al.* (1973). Twice as many girls as boys achieved reading scores at baseline and almost twice as many two years later. This is interesting in view of the initial significant difference in expressive language ability in favour of the girls and would suggest that their superiority in expressive language may have facilitated their acquisition of reading skills. By the end of the study only one girl in the mainstream group and one in the MLD group were non-readers; both were the least able and youngest. By contrast to the sex effect, no such apparent effect can be attributed to schooling, although there is a suggestion within the reading comprehension results that the MLD children show poorer performances than the mainstream children. Whilst twice as many mainstream children were reading at baseline, the difference was less marked by the end of the first year; by the end of the two-year period the mainstream children were only slightly better than the MLD children (77·8 per cent *vs.* 61·11 per cent). It is likely that, as in non-handicapped children, the sex difference overrides other factors contributory to reading development, such as educational environment (Rutter and Yule, 1975).

It is outside the remit of this study to attempt to make causal statements regarding the nature of the influence of mainstream education on the progress made by Down's Syndrome children. It would indeed be difficult to determine which aspects of the mainstream placements contributed to the improvements made by the children in just two years of study. However, the results, indicating as they do specific academic advantages in numeracy and comprehension in these mainstream children, as well as significant improvements in general cognitive ability and a trend in their favour on other measures, go some way to providing the sort of efficacy data for mainstreaming demanded by Zigler and Hall (1986). A wider ranging study is called for which would control more educational and social variables, e.g., class size, teacher support, teacher training, pupil support, social background, family size, etc. The evidence from this study is that Down's Syndrome children whose ability levels are congruent with those of children with moderate learning difficulties are capable of having their special educational needs met within the mainstream curriculum and the environment of the mainstream primary school.

ACKNOWLEDGMENTS.—The authors are grateful to Dr. A. R. Jonckheere for statistical advice and comments. Thanks are also due to all the children who took part in the study and to their parents and teachers.

Correspondence and requests for reprints should be addressed to Dr. Wendy Casey, Principal Clinical Psychologist, Harperbury Hospital, Radlett, Herts., WD7 9HQ.

## REFERENCES

BRADFIELD, H. R., BROWN, J., KAPLAN, P., RICKERT, E., and STANNARD, R. (1973). The special child in the regular classroom. *Except. Children*, 39, 384-390.

BRENNAN, W. K. (1982). *Special Education in Mainstream Schools: the Search for Quality*. London: National Council for Special Education

CARROLL, A. (1967). The effects of segregated and partially integrated school programs on self concept and academic achievement of educable mental retardates. *Except. Children*, 34, 93-99.

CORMAN, L., and GOTTLIEB, J. (1978). Mainstreaming mentally retarded children: a review of research *Int. Rev. Res. ment. Retard.*, 9, 251-257.

CRATTY, B. J. (1972). *Physical Expressions of Intelligence*. Englewood Cliffs, N.J.: Prentice-Hall.

CUNNINGHAM, C. C., GLENN, S. M., WILKINSON, P., and SLOPER, P. (1985). Mental ability, symbolic play and receptive and expressive language of young children with Down's Syndrome. *J. child Psychol. Psychiat.*, 26, 255-265.

GIBSON, D. (1978) *Down's Syndrome: the Psychology of Mongolism*. Cambridge: Cambridge University Press.

HEGARTY, S. (1982). Meeting special educational needs in the ordinary school, *Educ. Res.*, 24, 174-181.

HEGARTY, S., and POCKLINGTON, K (1981). *Educating Pupils with Special Needs in the Ordinary School.* Windsor: NFER-Nelson.

HEGARTY, S., POCKLINGTON, K., and LUCAS, D. (1982). *Integration in Action: Case Studies in the Integration of Pupils with Special Needs.* Windsor: NFER-Nelson.

HODGSON, A., CLUNIES-ROSS, L., and HEGARTY, S. (1984). *Learning Together: Teaching Children with Special Educational Needs in the Ordinary School.* Windsor: NFER-Nelson.

LORENZ, S. (1983). A tale of three children. *Parents' Voice*, 33 (3), 8-9.

LORENZ, S., SLOPER, T., and CUNNINGHAM, C C (1985). Reading and Down's Syndrome. *Br. J. spec. Educ.*, 12, 65-67.

OFFICE OF POPULATION CENSUSES AND SURVEYS (1970). *Classification of Occupations*. London: HMSO.

PIAGET, J. (1953) *The Origin of Intelligence in the Child.* London: Routledge.

REBELSKY, F. G., STARR, R. H., and LURIA, Z. (1967). Language development. the first four years. In BRACKBILL, Y (Ed.), *Infancy and Early Childhood*. New York: The Free Press.

SATTLER, J. M. (1982). *Assessment of Children's Intelligence and Special Abilities.* Boston: Allyn and Bacon.

STOBART, G. (1986). Is integrating the handicapped psychologically defensible? *Bull. Br. psychol. Soc.*, 39, 1-3.

WALL, W. D. (1979). *Constructive Education for Special Groups: Handicapped and Deviant Children.* London: Harrap.

WARNOCK REPORT (1978). *Special Educational Needs* Report of the Committee of Inquiry into Education of Handicapped Children and Young People. London: HMSO.

YOSHIDA, R. (1986). Setting goals for mainstream programs. In MEISEL, C. J. (Ed.), *Mainstreaming Handicapped Children: Outcomes, Controversies, and New Directions.* Hillside, N J.: Lawrence Erlbaum.

ZIGLER, E., and HALL, N. (1986). Mainstreaming and the philosophy of normalisation. In MEISEL, C. J. (Ed.), *Mainstreaming Handicapped Children: Outcomes, Controversies, and New Directions.* Hillside, N.J.: Lawrence Erlbaum.

(*Manuscript received* 10*th March*, 1987)

*Br. J. educ. Psychol.*, 58, 287-300, 1988

# TEXT SEGMENTING AND COMPREHENSION: EFFECTS OF READING AND INFORMATION PROCESSING ABILITIES

BY JOHN R. KIRBY
*(Queen's University, Kingston, Ontario, Canada)*

AND CHRISTOPHER J. GORDON
*(New South Wales Department of Education, Bathurst, Australia)*

SUMMARY. A study is reported which tests the hypothesis that segmenting text into appropriate syntactic units will facilitate the comprehension of readers who have adequate word-level skills but poor comprehension skills. This specific hypothesis is not confirmed, but evidence is presented that poor comprehenders, regardless of word-level skills, do benefit from text segmenting. A second hypothesis tested concerns the information processing skills involved in syntactic analysis. Results confirm that simultaneous processing is involved in the forming of syntactic units. In conditions in which the text is already segmented, however, successive processing is shown to be involved, as well as simultaneous processing, it is argued, at a higher level. These results suggest that syntactic analysis is a potential bottleneck in reading, and that training studies should be designed, employing simultaneous and successive processing, to improve syntactic analysis skills and facilitate reading comprehension.

## INTRODUCTION

SEVERAL distinct levels of text processing can be identified in reading. One of these is certainly the word identification level, and earlier ones include visual and phonological analysis. Beyond the word level however, other levels of processing concern syntactic analysis and the extraction of meaning, the latter being subdivided further with respect to the degree of abstraction.

The syntactic analysis level has received relatively little attention, although it appears to serve a critical function in the processing of text information. Its function is critical because of working memory limitations. If a reader had to hold in working memory all of the words of a sentence before meaning could be processed, then that number of words would in most cases exceed working memory capacity and comprehension would fail. Syntactic analysis allows the reader to segment the input information into chunks, a smaller number of higher order units, roughly corresponding to phrases or clauses. These units are in general smaller than the micropropositions of Kintsch and van Dijk (1978) or the idea units of other investigators (e.g., Mandler and Johnson, 1977). Syntactic analysis allows the reader to transform a 10- or 15-word sentence into three or four chunks, a number well within working memory limits and which allows comprehension to proceed. The purpose of this paper is to explore further the syntactic analysis or text segmenting stage, to investigate the effects of assisting subjects with syntactic analysis, and to explore the information processing abilities which may underlie the text segmenting process.

Theory would predict a relationship between reading ability and the ability to segment text for two reasons. In the first place, text segmenting ability would seem to be a prerequisite for skilled reading: if readers are unable to smoothly or automatically process syntactic units in text, it is unlikely that they will be able to progress to deeper levels of meaning extraction. This reason relates to a model in which lower-level psycholinguistic processes affect higher-level ones; Perfetti and Lesgold (1979), for example, describe a comparable explanation in their bottleneck hypothesis, in which inefficient word-encoding inhibits higher-level reading processes. The second reason for a relationship between segmenting ability and reading is top-down in nature: greater comprehension of a passage in general terms would allow sentence-level expectations to guide syntactic analysis and ensure more efficient encoding of syntactic units. Thus syntactic analysis may be a prerequisite for skilled reading as well as a by-product of skilled reading.

One method of investigating syntactic analysis skills is to manipulate the structure of texts so that their syntactic structure is more apparent (Cromer, 1970; Frase and Schwartz, 1979; Mason and Kendall, 1979; O'Shea and Sindelar, 1983). For example, Cromer (1970) identified a group of readers, whom he termed "difference" readers, whose reading comprehension scores increased after reading text which had been syntactically segmented. Surprisingly, this result was found with university students. Cromer suggested that these had adequate word level or vocabulary skills, but approached the reading task in such a way that they did not combine words into syntactic chunks and therefore proceed adequately towards meaning. Thus, Cromer conceived that the effect of text segmenting was to facilitate the extraction of meaning for these students. This result was supported by Frase and Schwartz (1979), who found that segmenting decreased adults' reading times. Subsequent studies with younger readers have yielded more ambiguous results. Mason and Kendall (1979) found that segmented text increased reading times and improved reading comprehension scores for less able fourth graders, but had no effect upon high ability or middle ability readers. In contrast, however, O'Shea and Sindelar (1983) found that text segmenting yielded higher comprehension scores in both good and poor readers.

In spite of the results of O'Shea and Sindelar (1983), theory would predict that text segmenting should assist less able rather than more able readers, the latter group being presumed to engage already in that activity. This, however, must depend upon text difficulty, in that a good reader may be able to analyse syntactically a text at his or her reading level, but not be able to do so with a more difficult text. Further difficulties may arise because of subjects' different interpretations of what is required in the task, and how comprehension is measured. One reason why the Mason and Kendall (1979) and O'Shea and Sindelar (1983) studies may have failed to find consistent results is that they did not specifically take into account their pupils' vocabulary skills. As Cromer (1970) argued, the effect of text segmenting should be specific to that group whose vocabulary skills are high but whose comprehension skills are low.

The study to be reported in this paper makes use of three text conditions: *Appropriate Segments*, in which the text segments correspond to syntactic units; *Inappropriate Segments*, in which the text units do not correspond to syntactic units, and *Single Words*, in which the text is presented word by word. If Cromer's hypothesis is correct, those subjects whose vocabulary scores are high but whose comprehension scores are low should benefit most from the Appropriate Segments condition.

For several reasons the approach to text segmenting described in the preceding section is not completely satisfactory. In the first place, reading comprehension and reading vocabulary abilities are relatively gross measures of individual differences. Furthermore, while it is possible to separate vocabulary and comprehension abilities theoretically, in practice they are closely linked. Vocabulary knowledge is necessary for comprehension, and comprehension skills can be used to circumvent missing elements of vocabulary knowledge. This is supported by the high correlation between vocabulary and comprehension scores on most reading tests. A different approach to examining text segmenting is to consider the information processing skills that are involved in it. Such an analysis would provide another dimension to the examination of text segmenting and may provide the basis for identifying problems readers through non-reading measures.

One model of information processing abilities that has proved useful in examining reading problems has been described by Das *et al.* (1975, 1979; see also Cummins and Das, 1977; Kirby and Das, 1977, 1978; Robinson and Kirby, in press). This model is based upon Luria's (1966) neuropsychological theory of information integration and describes two information processing abilities, simultaneous processing and successive processing. Simultaneous processing is the ability to take a set of information elements and integrate them into a unitary representation. Successive processing is the complementary ability to take a set of information elements and form them into a temporal sequence. Both types of processing are involved in most tasks; however many tasks depend primarily upon one or the other ability. Das *et al.* (1979) have identified a battery of tests which can be used to define simultaneous and successive processing abilities. Psychometrically, simultaneous and successive processing represent clusters of more narrow abilities which share common processing characteristics. In this way they would be analogous, for example, to verbal and performance scores on the WISC-R. Other test batteries exist to measure simultaneous and successive processing (e.g., Kaufman and Kaufman, 1983).

In reading, successive processing is related to the phonological decoding of individual words and to the maintaining of word order, in order to allow syntactic analysis. Simultaneous processing is related to the immediate recognition of whole words and to the forming of text units, be they at the syntactic phrase or propositional levels (Cummins and Das, 1978; Kirby, 1982; Robinson and Kirby, in press). The segmenting of text into syntactic units is an interesting example of a skill which relies upon both simultaneous and successive processing abilities. Successive processing is required to retain the words in correct syntactic order, so that simultaneous processing can be employed to form the appropriate syntactic chunk. In general, as the number of words in the syntactic chunk is usually well within the working memory limitations of average readers, it is likely that successive processing would be less related to syntactic chunking (and therefore reading comprehension) than would simultaneous processing.

Thus one would predict that if text segmenting ability were being measured, children with high simultaneous processing ability would do better than children with low simultaneous processing ability, while children of high and low successive processing abilities should not differ to the same degree. This situation should arise in the Single Words condition (described earlier), in which subjects have to do their own text segmenting.

However, what should happen when the text is segmented for the readers? The Appropriate Segments condition should facilitate the reading comprehension performance of those children with low simultaneous processing ability, because less emphasis is placed upon their area of weakness. In the Inappropriate condition, however, children with low simultaneous processing ability should be penalised, as in the Single Words condition, this time because they must overcome an inappropriate organisation of the text.

## METHOD

### *Sample*

The pupils selected for this experiment were 352 primary school children from grades 3, 4, 5 and 6 of a middle class school in Newcastle, Australia. The children were distributed across grades as follows: 88 in grade 3, 80 in grade 4, 95 in grade 5 and 89 in grade 6. Each grade had an approximately equal balance of males and females.

### *Tests*

Tests were administered to measure reading vocabulary, reading comprehension, simultaneous and successive processing, and performance under altered text conditions.

*Vocabulary.* The test of reading vocabulary was the Primary Reading Survey test of Word Knowledge (ACER, 1976). Appropriate levels of this test were employed for each of the grade groups. This test consists of multiple choice synonym items, in which words are presented without any context, thus minimising the effects of comprehension. Scores on this test were used only to divide pupils in each grade into three groups, those low, middle or high in reading vocabulary knowledge relative to their grade.

*Simultaneous processing.* Two tests were used to measure simultaneous processing. The Coloured Progressive Matrices (Raven, 1956) has been used frequently before as a measure of simultaneous processing (Das *et al.*, 1979). This test has 36 items, each consisting of a design with a piece missing. The child's task is to select from a set of six possible alternatives the correct piece which completes the pattern. This test is a measure of simultaneous processing because it is necessary for the child to hold mentally the entire design while determining the missing piece. The second measure of simultaneous processing was Figure Copying (Das *et al.*, 1979). This test requires subjects to inspect a geometric figure and then reproduce it by drawing. Reproductions are scored for their maintenance of geometric relations rather than for strict accuracy. Further details of these two tests can be found in Das *et al.* (1979).

*Successive processing.* Two measures of successive processing were also employed. The first of these, Digit Span Forward, has been used frequently in the past (Das *et al.*, 1979). In this test, individuals are presented with a series of numbers spoken at the rate of one per second, and are then required to repeat the series in the same order as it was presented. Digit series increase in length until the pupil fails two trials at the same length, at which point testing is discontinued. The child's score is the greatest list length successfully recalled.

The second measure of successive processing was designed for this study, partly to overcome criticisms that measures of successive processing assess only rote memory. In the Sentence Repetition test, individuals are presented orally with sentences which increase in length and are then asked to repeat them. Recall is scored not for the number of words correct but rather for the number of syntactic chunks correctly recalled. For this study, sentences ranged in length from two chunks (e.g., "The paint dribbled/down the brush") to seven chunks (e.g., "John had the trumpet/on a chair/by his bed/and throughout the night/he kept picking it up/and going through the motions/of playing"). Testing proceeds until an individual fails twice to recall successfully sentences of a given chunk length. Each individual's score is the length, in chunks, of the longest sentence correctly recalled. This sentence repetition test can be seen to test more than rote memory because the readers must process the syntax, and perhaps some semantic meaning, of the sentences in order to chunk them appropriately. Further details about this test may be found in Gordon (1982).

*Comprehension of normal and altered text.* Reading comprehension was measured in conjunction with performance in the altered text conditions. Reading passages selected from the Progressive Achievement Test (ACER, 1970) were used for these purposes. Two reading passages were selected at each of the grade 3, 4, 5 and 6 reading levels, one from each of the two parallel versions of the test. Each of these passages is normally followed by five multiple choice comprehension items. In order to increase the range of possible scores, five additional multiple choice items were constructed for each reading passage. These additional items were modelled upon the five original items. While no attempt was made to equate the original and additional items psychometrically, differences between old and new items would be unlikely to affect the results as all children within a grade group were to answer all items.

Four different versions of each reading passage were constructed. The Normal Text version was presented in the same format as a regular reading passage, typed on a sheet of paper and with no special indentation or spacing. The Single Words version of each passage was prepared with one word per line in a vertical column down the page, a passage consisting of five to seven columns. Therefore children had to read down a column of words to the bottom of the page and then go back to the top of the page to read the second column and so on. No special spacing was employed between phrases or sentences. In the Inappropriate condition the passages were typed so that the words were grouped in chunks which did not correspond to appropriate syntactic units. In the Appropriate condition, a similar format was employed with the exception that chunks corresponded to appropriate syntactic units. For these purposes, a syntactic unit was defined as a phrase node that did not contain an embedded phrase node; non-phrase nodes, or those which did not branch, were attached to the following phrase node (see Gordon, 1982, for further details). For both the Appropriate and Inappropriate conditions, chunks within a sentence were progressively indented across the page, new sentences beginning again on the left margin. See Table 1 for examples of both conditions. In all conditions, normal punctuation was maintained and sentences began with a capital letter.

TABLE 1

EXAMPLES OF PARTS OF TEXTS PRESENTED IN THE INAPPROPRIATE AND APPROPRIATE SEGMENTS CONDITIONS

*Inappropriate Segments*

Michael was born in an industrial
city in Northern
England where he spent the
early years of
his life.

*Appropriate Segments*

Michael was born
in an industrial city
in Northern England
where he spent the early years
of his life.

In order to equate the difficulty of the passages employed at the different grade levels, scores were standardised for each passage within each grade level, such that the overall mean across Normal, Single Words, Inappropriate and Appropriate conditions was 0·0 with a standard deviation of 1·0. A similar procedure was followed with vocabulary scores. Thus the vocabulary and comprehension measures employed here are relative to a subject's grade grouping. These procedures were followed in order to eliminate spurious grade effects due to different materials, while allowing for the possibility of interactions between grade and other factors.

### *Procedure*

Pupils were tested in three sessions, two in class groups and one individually, within a period of not more than three weeks. Both group sessions were of 40 minutes duration and were in the mornings of different days. Half of the pupils received the Coloured Progressive Matrices and Figure Copying in the first session, while the others received the Comprehension and Vocabulary measures in the first session. The individual tests, Digit Span and Sentence Repetition, were administered at convenient times throughout the school day.

All pupils received one passage in the Normal Text condition and one in one of the three altered text conditions. Half of them received the normal text first, half received the altered text first. Passages and passage order were counterbalanced across all subjects. Assignment of children to experimental conditions was done on a random basis.

## RESULTS

The results of this study were analysed by means of a series of analyses of variance in which the independent variables were grade, vocabulary ability, comprehension ability (from the Normal Text condition), simultaneous and successive processing abilities and text condition. Separate analyses were planned to investigate various combinations of these factors. In each the dependent variable was comprehension performance in the altered text conditions.

The first analysis performed was a 4 (Grade) × 3 (Comprehension Group) × 3 (Text Condition) analysis of variance. The purpose of this analysis was to investigate the possibility that children at different grade levels were differentially affected by the three text conditions. A significant main effects for grade was not possible, due to the standardisation across grades; furthermore, there was no significant interaction involving grade. Therefore grade was not included in the subsequent analyses as a factor.

### *Reading ability analyses*

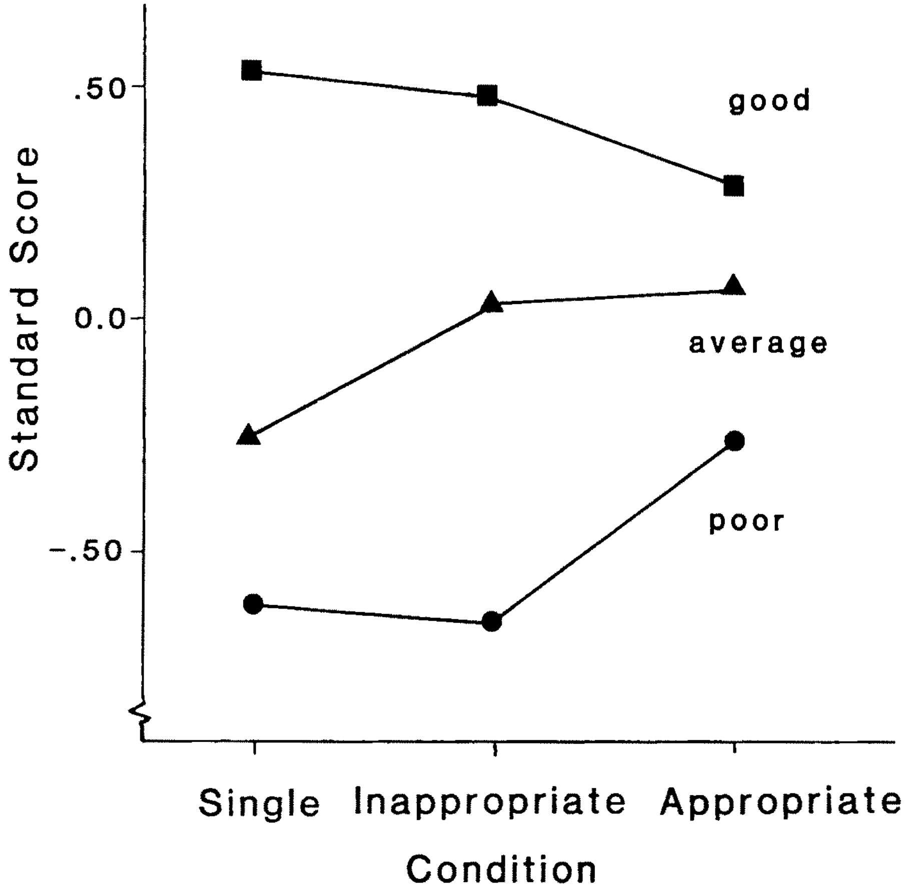
The first analysis to be reported was a 3 (Comprehension Group) × 3 (Vocabulary Group) × 3 (Text Condition) analysis of variance. The comprehension and vocabulary scores were divided into three groups (at the scores of $-0{\cdot}4$ and $+0{\cdot}4$) because Cromer's predictions seem most relevant for pupils in the bottom third of the comprehension distribution. The poor, average and good vocabulary groups consisted respectively of 128, 104 and 120 children; the poor, average, and good comprehension groups consisted of 109, 104, and 139.

This analysis yielded two significant main effects and one interaction. The comprehension group main effect, $F(2,325) = 8{\cdot}40$, $P < 0{\cdot}001$, indicated that performance increased with comprehension ability. The vocabulary group main effect, $F(2,325) = 22{\cdot}46$, $P < 0{\cdot}001$, indicated that performance also increased with vocabulary ability. The greater strength of the vocabulary effect is most likely due to the fact that vocabulary was measured with a full vocabulary test whereas comprehension ability was measured only by performance after one reading passage.

The only significant interaction, between Comprehension Group and Text Condition, $F(4,325) = 2{\cdot}79$, $P < 0{\cdot}05$, is illustrated in Figure 1. Tukey tests (at the 0·05 level) were applied to investigate group differences. In the Single Words

FIGURE 1

PERFORMANCE OF THREE GROUPS OF READERS IN THREE CONDITIONS



condition, good comprehenders performed significantly better than both average and poor comprehenders, who did not differ. In the Inappropriate condition good and average comprehenders did not differ, but both performed significantly better than poor comprehenders. In the Appropriate condition, the only significant difference was between good comprehenders and poor comprehenders. Comparing text conditions within ability groups, the only significant difference was between the poor comprehenders in the Appropriate and Inappropriate conditions.

There is no evidence in these results to support Cromer's contention that the Appropriate condition has maximum benefit for those above average in vocabulary ability but below average in comprehension. However, it does appear that the different text conditions have different effects upon the three comprehension groups. The groups appear most distinct in the Inappropriate condition, the condition which is most similar to normal reading. In contrast, the Single Words condition inhibits the performance of the average comprehenders, in that they perform no better than the poor comprehenders in that condition. More interesting is the fact that the groups are most similar in the Appropriate condition, primarily because of the significant boost this condition has upon the performance of the poor comprehenders. While Cromer's specific hypothesis is not confirmed, the more general hypothesis that text segmenting should benefit poorer readers is supported.

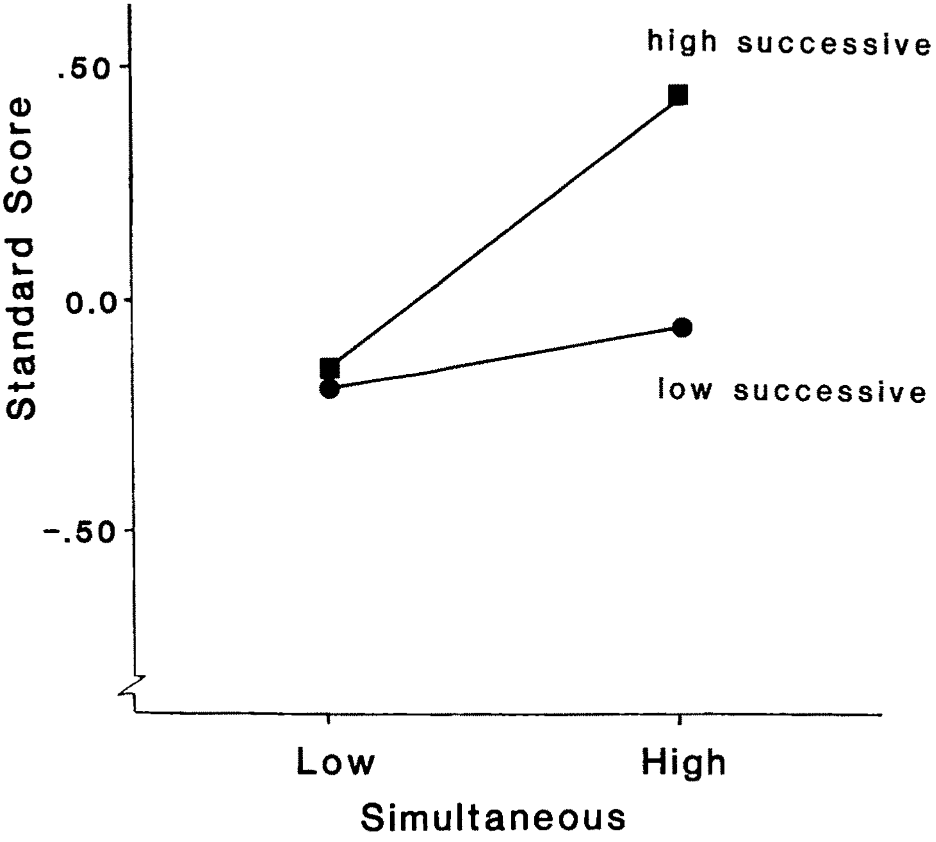
*Information processing analyses*

The scores for the four simultaneous and successive processing tests were correlated and submitted to a principal components analysis. The results of this analysis conformed to those of previous studies (Das *et al.*, 1979). Two factors had eigenvalues greater than 1·0, and when rotated to a varimax criterion these factors clearly defined simultaneous and successive processing factors. These factors were then used to generate factor scores for each pupil, and pupils were divided at the factor score medians to form four groups. Because 18 pupils had not completed these measures, there were 85 in the low simultaneous-low successive group, 82 in the low simultaneous-high successive, 82 in the high simultaneous-low successive, and 85 in the high simultaneous-high successive.

A Vocabulary × Comprehension × Simultaneous × Successive × Text Condition analysis of variance was attempted, but could not be performed due to empty cells. Because of the importance of the Comprehension factor in the previous

FIGURE 2

PERFORMANCE OF SUBJECTS ACCORDING TO LEVELS OF SIMULTANEOUS AND SUCCESSIVE PROCESSING



analysis, it was decided to perform a 3 (Comprehension Group) × 2 (Simultaneous processing) × 2 (Successive processing) × 3 (Text Condition) analysis of variance.

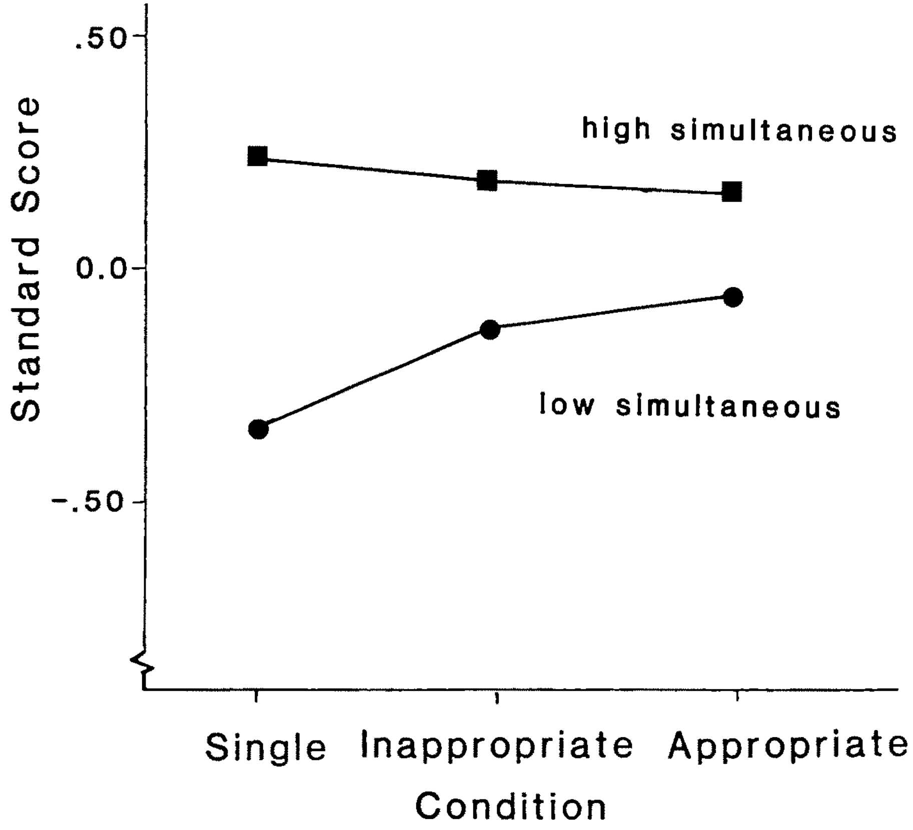
Our prediction regarding this analysis was that performance in the text conditions would depend upon simultaneous processing abilities: greater simultaneous processing skill should be associated with greater performance under the Single Words and Inappropriate Segments conditions, and the difference between high and low simultaneous processing subjects should decrease in the Appropriate Segments condition. Although this overall pattern did emerge, the difference in the Inappropriate condition was not significant, and the results were dependent as well upon pupils' successive processing scores.

The significant Simultaneous, F (1,292) = 9·61, P < 0·005, Successive, F (1,292) = 7·72, P < 0·01, and Comprehension, F (2,292) = 25·03, P < 0·001, effects all indicated, as expected, that the more able pupils performed better across text conditions.

Also reliable was the Simultaneous × Successive effect, illustrated in Figure 2. This interaction is clearly due to the superior performance of the group above average in both information processing skills, while the other three groups do not differ from each other. This effect is quite consistent with the suggestion from

FIGURE 3

PERFORMANCE OF HIGH AND LOW SIMULTANEOUS PROCESSING SUBJECTS IN THREE CONDITIONS

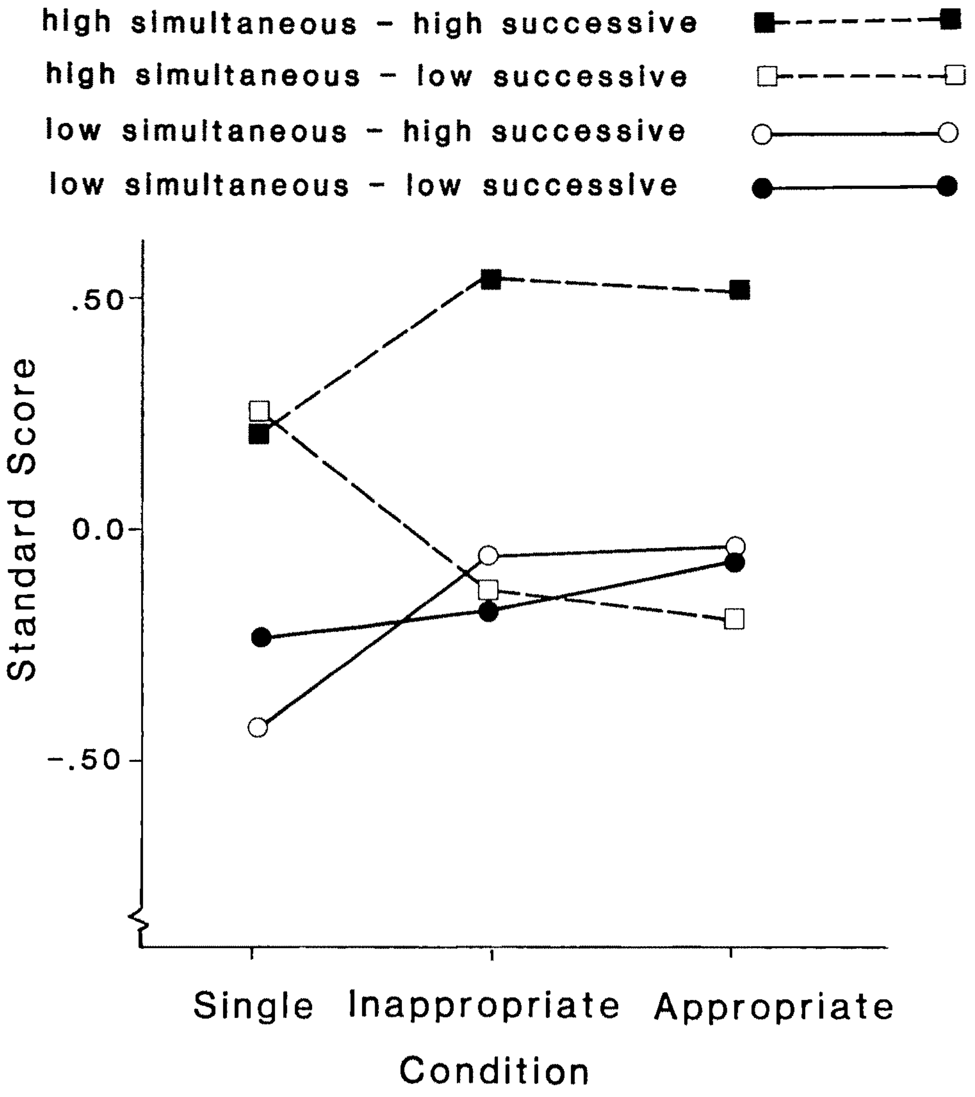


previous literature that both simultaneous and successive processing abilities are required for successful reading comprehension performance (e.g., Kirby and Das, 1977; Kirby, 1982).

The second interaction, Simultaneous × Successive × Text Condition, F (2,298) = 3·49, $P < 0{\cdot}02$, contains the predicted effect concerning Simultaneous processing and Text Condition. While the latter effect was not significant, it merits investigation due to its theoretical importance, and so is illustrated in Figure 3. In this figure it can be seen that the two Simultaneous processing groups are most

FIGURE 4

PERFORMANCE OF SUBJECTS ACCORDING TO LEVELS OF SIMULTANEOUS AND SUCCESSIVE PROCESSING, IN THREE CONDITIONS



different in the Single condition ($P < 0{\cdot}01$, Tukey test) and least different in the Appropriate condition ($P > 0{\cdot}05$). Somewhat surprising is the only moderate difference between groups in the Inappropriate condition (these two groups are different by 0·32, and the Tukey test critical difference is 0·35).

This trend, however, is affected by the students' successive processing scores; thus the three-way interaction. It can be seen in Figure 4 that pupils of high simultaneous processing ability perform quite differently in the Inappropriate and Appropriate Segments conditions, depending upon their level of successive processing. Those with high successive processing scores do well in both chunked conditions, whereas those with low successive processing scores do no better than pupils with low simultaneous processing scores in these conditions. Performance of the low simultaneous processing pupils, and of all in the Single Words condition, is not affected by successive processing.

The predicted dependence of performance in the text conditions upon simultaneous processing is apparent in these results, but is not as strong as expected, particularly in the Inappropriate condition. Furthermore, the effect is obscured by the involvement of successive processing. These findings, and the surprising lack of difference between the Inappropriate and Appropriate conditions, merit further discussion.

## DISCUSSION

Two sets of hypotheses regarding the effect of text segmenting were investigated in this study, one with respect to reading skills and the other with respect to information processing skills. In each case the hypotheses were confirmed in broad terms, but were shown to be in need of revision concerning detail or the involvement of other factors.

### *Text segmenting and reading abilities*

The major hypothesis tested in the reading skills analysis was derived from Cromer's (1970) research, that the Appropriate Segments condition should most benefit those readers with good word-level (i.e., vocabulary) skills but poor comprehension skills. This hypothesis ran counter to other studies which had shown text segmenting to benefit either broadly defined poor readers (Mason and Kendall, 1979), or all readers (O'Shea and Sindelar, 1983).

The results of the present study failed to support Cromer's specific hypothesis, but did show that poor comprehenders, regardless of vocabulary skills, benefited from the text segmenting. The large number of pupils and the range of grades employed in this study provide a generous test of the original hypothesis; accordingly it must be concluded that it is untenable for elementary school children. It may of course remain valid for the college population who provided Cromer's (1970) subjects.

The benefit of the Appropriate Segments condition for the poor comprehenders confirms the findings of Mason and Kendall (1979), and suggests that the syntactic analysis stage in processing is acting as a "bottleneck" for these readers. If this is the case, then instruction at the level of sentence parsing may be very helpful for children with poor comprehension skills. Furthermore, children with a difficulty at the syntactic analysis level may be less able to benefit from instruction aimed at higher comprehension levels, for instance at the extraction of main ideas from larger texts. Top-down factors may assist these children to some extent, but there seems little reason for avoiding instruction in the syntactic skills in which they are weak.

Somewhat surprising in these results is the lack of difference between the Inappropriate and Appropriate conditions for all but the least able comprehenders.

There is no suggestion in the data that this is due to a ceiling effect. Given the intuition of most adults and children that the Appropriate texts are easier to read, which is supported by various studies with respect to comprehension test performance and reading time (Cromer, 1970; Frase and Schwartz, 1979; Mason and Kendall, 1979; O'Shea and Sindelar, 1983), this suggests that most readers in the Inappropriate condition in the present study were able to adapt their reading to minimise the effects of the inappropriate segmenting.

Perhaps this should not be surprising, in that texts are normally formatted in an inappropriate segments manner with little regard to how lines end and begin. In fact, because all of the lines in the Inappropriate condition consisted of two half-phrases, it could be that the more able students simply learned to read from the middle of one line to the middle of the next; by in effect giving a cue as to where the next syntactic phrase break was to occur, the Inappropriate Segments condition may have become easier than a normal reading task. It is possible that controlling the presentation of the text, so that readers could only see one segment at a time, would increase the difference between the Inappropriate and Appropriate conditions. From the present results, however, it is worth concluding that many readers are able to adapt their reading processes to overcome experimental conditions.

Can text segmenting ever benefit better readers? There is no evidence in the present study that it can, but this conclusion is limited by the pupils, reading materials, and type of segments employed. Other researchers have shown benefits for adults (Cromer, 1970; Frase and Schwartz, 1979) who were surely more able readers than those included in the present study. Two possibilities can be suggested here: that better readers may only be assisted with much more difficult reading materials (cf. Mason and Kendall, 1979, p. 75), and that they may require segmenting at a higher level of analysis than the syntactic phrase level.

### *Text segmenting and information processing*

Das *et al.*'s (1979) theory of simultaneous-successive processing was used to derive the hypothesis that simultaneous processing skills should determine performance in the Single Words and Inappropriate conditions, while the Appropriate condition should allow subjects with low simultaneous processing skills to overcome their weakness and perform better. In general terms this pattern of results was found, though the difference between the high and low simultaneous processing groups in the inappropriate condition failed to attain significance. This finding may be explained by the argument provided in the preceding section, that the Inappropriate condition inadvertently contained cues about the location of breaks between syntactic units. Cues concerning syntactic analysis would serve to lessen the need for simultaneous processing at this level, accounting for the result that simultaneous processing group differences in this condition were midway between those in the Single and Appropriate conditions.

Of more importance theoretically is the evidence for the involvement of successive processing in the two segmented conditions, for the subjects of greater simultaneous processing ability. One possible interpretation of this is based upon the suggestion that the simultaneous processing required in the Single Words and segmented conditions is at two different levels. Whereas simultaneous processing is required in the Single Words condition to form syntactic units, the two segmented conditions already provide this level of processing. They instead demand successive processing for the retention of two or more syntactic units, and then simultaneous processing to form higher-level units of meaning from these phrase units, perhaps at the sentence level. Thus, only readers with good levels of both types of processing skills would be able to capitalise upon the segmented conditions to maximise their performance (as is shown in Figure 4).

These interpretations, particularly that concerning the involvement of successive processing, are somewhat speculative and *post hoc*, but are consistent with the Das *et al.* (1979) theory. The confirmation of the hypothesis regarding simultaneous processing in the text conditions is experimental support for a theory whose major support has been factor analytic in nature. The implication of successive processing, and of simultaneous processing of a higher level, in the segmented text conditions is consistent with a cyclical description of information processing through a hierarchy of levels (Kirby and Biggs, 1981). Simultaneous processing at a low level is used to identify units (e.g., letters) which are then held in order by means of successive processing to allow simultaneous processing to form a higher level unit (a word). Such a cycle would alternate between simultaneous and successive processing through the levels of, for instance, letters, words, syntactic phrases, idea units, main ideas, and themes (Kirby, 1988; Kirby and Williams, in preparation, chapter 7).

### *Conclusions*

The results of this study have underlined the importance of syntactic analysis in reading, and shown it to be a potential processing bottleneck for poor readers. As such it merits further study, in particular to see whether children's syntactic analysis skills can be improved through training, and then whether those improvements lead to increased reading comprehension.

The technique of text segmenting also deserves further study, both as a research tool, and as an instructional device. While the present study did not show any benefit of segmenting for the more able readers in grades 3 to 6, the comprehension of the less able readers was assisted. Presumably segmenting would also be of help to younger readers, guiding them toward correct sentence parsing and toward meaningful reading.

This study also explored the information processing skills underlying text segmenting, confirming the hypothesis that simultaneous processing was primarily involved. More unexpected was evidence for a cycle of simultaneous and successive processing, leading to higher levels of comprehension. These information processing skills could form the basis of the training programme referred to above: simultaneous processing would have to be addressed to improve text segmenting skill, and successive processing would have to be involved if that skill is to contribute to increased comprehension.

Correspondence and requests for reprints should be addressed to Dr. J. R. Kirby, Faculty of Education, Queen's University, Kingston, Ontario, Canada, K7L 3NG.

## REFERENCES

Australian Council for Educational Research (1970). *Progressive Achievement Tests. Reading Comprehension.* Hawthorn, Victoria, Australia: ACER.

Australian Council for Educational Research (1976). *Primary Reading Survey Tests.* Hawthorn, Victoria: ACER.

Cromer, W. (1970). The difference model: a new explanation for some reading difficulties. *J. educ. Psychol.*, 61, 471-483.

Cummins, J., and Das, J. P. (1977). Cognitive processing and reading difficulties: a framework for research. *Alberta J. educ. Res.*, 23, 245-256.

Cummins, J., and Das, J. P. (1978). Simultaneous and successive syntheses and linguistic processes. *Int. J. Psychol.*, 13, 129-138.

Das, J. P., Kirby, J., and Jarman, R. F. (1975). Simultaneous and successive syntheses: an alternative model for cognitive abilities. *Psychol. Bull.*, 82, 87-103.

Das, J. P., Kirby, J. R., and Jarman, R. F. (1979). *Simultaneous and Successive Cognitive Processes.* New York: Academic Press.

Frase, L. T., and Schwartz, B. J. (1979). Typographical cues that facilitate comprehension. *J. educ. Psychol.*, 71, 197-206.

GORDON, C. J. (1982). A comparative study of the role of word-based skills and processing strategies in the reading comprehension of primary school children. Unpublished Master of Psychology (Educational) thesis, University of Newcastle, N.S.W.

KAUFMAN, A. S., and KAUFMAN, N. L. (1983). *Kaufman Assessment Battery for Children. Interpretive Manual.* Circle Pines, Minn.: American Guidance Service.

KINTSCH, W., and VAN DIJK, T. A. (1978). Toward a model of text comprehension and production. *Psychol. Rev.*, 85, 363-394.

KIRBY, J. R. (1982). Cognitive processes, school achievement, and comprehension of ambiguous sentences. *J. psycholinguistic Res.*, 11, 485-499.

KIRBY, J. R. (1988). Style, strategy, and skill in reading. In SCHMECK, R. R. (Ed.), *Learning Styles and Learning Strategies.* New York: Plenum Press.

KIRBY, J. R., and BIGGS, J. B. (1981). *Learning Styles, Information Processing Abilities, and Academic Achievement.* Final report to the Australian Research Grants Committee. University of Newcastle, N.S.W.

KIRBY, J. R., and DAS, J. P. (1977). Reading achievement, IQ, and simultaneous-successive processing. *J. educ. Psychol.*, 69, 564-570.

KIRBY, J. R., and DAS, J. P. (1978). Information processing and human abilities. *J. educ. Psychol.*, 70, 58-66.

KIRBY, J. R., and WILLIAMS, N. H. Learning Problems: An Integrated Approach. (In preparation.)

LURIA, A. R. (1966). *Human Brain and Psychological Processes.* New York: Harper and Row.

MANDLER, J. M., and JOHNSON, N. S. (1977). Remembrance of things parsed: story structure and recall. *Cogn. Psychol.*, 9, 111-151.

MASON, J. M., and KENDALL, J. R. (1979). Facilitating reading comprehension through text structure manipulation. *Alberta J. educ. Res.*, 25, 68-76.

O'SHEA, L. J., and SINDELAR, P. T. (1983). The effects of segmenting written discourse on the reading performance of low- and high-performance readers. *Reading Res. Q.*, 18, 458-465.

PERFETTI, C. A., and LESGOLD, A. M. (1979). Coding and comprehension skill in reading and implications for reading instruction. In RESNICK, L. B., and WEAVER, P. A. (Eds.), *Theory and Practice of Early Reading. Vol. 1.* Hillsdale, N.J.: Erlbaum.

RAVEN, J. C. (1956). *Coloured Progressive Matrices.* London: H. K. Lewis.

ROBINSON, G. L. W., and KIRBY, J. R. (in press). An information processing approach to language and reading problems. *Aust. J. hum. Commun. Disorders.*

(*Manuscript received* 20*th July*, 1987)

*Br. J. educ. Psychol.*, 58, 301-306, 1988

# DETECTING FALSE STATEMENTS IN TEXT: THE ROLE OF OUTLINES AND INSERTED HEADINGS

By JOHN A. GLOVER AND DAMON KRUG
*(Ball State University, Indiana, USA)*

SUMMARY. Students read text material containing false statements under four instructional conditions. Given directions to underline anything "hard to understand", students in a control condition identified only about 26 per cent of the false statements. Given the same directions, students who carefully read an outline of the text or who carefully read text headings as they proceeded through the material identified approximately 56 per cent of the false statements. Finally, in a condition in which students read an outline prefacing the text, and also carefully read inserted headings, they identified approximately 87 per cent of the false statements. The results are discussed in terms of schema theory.

## INTRODUCTION

THE inability of many students to make correct judgments about the accuracy of information they read has been of concern to psychologists and educators since early in this century (see Huey, 1908) and remains an important issue among contemporary educators and psychologists attempting to understand the processes involved in reading comprehension (e.g., Harris and Sipey, 1983; Baker and Brown, 1984, Foorman and Siegel, 1985; Baker and Wagner, 1987; Just and Carpenter, 1987). Indeed, several recent studies provide clear evidence that readers from elementary school through college-age typically do not monitor their understanding of prose on the basis of criteria for critical reading (Baker and Anderson, 1982; Glenberg *et al.,* 1982; Baker, 1984, 1985; Epstein *et al.*, 1984; Baker and Wagner, 1987).

The importance of critically evaluating what is read seems clear. Students should be able to compare and contrast various arguments (e.g., two different views of the Watergate scandal) and detect potential discrepancies that may appear in various writings (e.g., some political essays, "slanted" editorials). One facet of critical reading is the ability of students to identify statements that contradict their knowledge (Glenberg *et al.*, 1982; Baker and Brown, 1984). Consider the following example taken from Baker and Wagner (1987, p. 248):

> Governors frequently go to great lengths to win legislators over to their side. They often spend many hours simply socialising with them. As governor of Montana, Ronald Reagan used to invite groups of legislators to his home. There, among other things, they would play with the model electric train network that he had set up in his basement.

In Baker's (1985) study, only 24 per cent of college students asked to read a set of passages and underline anything hard to understand or incorrect identified the incorrect statement about Montana. Further, only 51 per cent who were told that there would be inconsistencies with prior knowledge in the passages identified the incorrect statement. Interestingly, when Baker later questioned the students, she found that the vast majority knew that Reagan had been governor of California and not Montana. In our own pilot work, only one of 24 students who had read *The Lord of the Rings* identified the statement, "Bilbo, who had once served Sauron, loved Rivendell", as containing an inaccuracy. Similarly, only six of 30 students recruited from a college-level European history course found anything odd in the statement, "Hitler, whose American mother left Baltimore in 1885, served in World War I". Upon debriefing, all of the students who had read about Bilbo knew that he never served Sauron. Further, all of the history students were absolutely certain that Hitler's mother was not American.

To this point, the evidence indicating that readers do not typically identify elements of text that disagree with their knowledge seems very clear. What is not yet clear are procedures whereby students' identification of false statements can be facilitated. (Note: Realising that "false" is a relative term dependent on the information in people's knowledge bases, it is nonetheless a much more convenient term to employ in the remainder of this article than, say, "knowledge based-incongruent" statements.) However, a review of the literature on techniques designed to influence readers' comprehension of text indicates several methods that may hold promise. In particular, the use of outlines prefacing text, and headings inserted in text, may be procedures that affect students' ability to identify false statements.

Outlines depict the topics within a text, and the order in which the topics occur. They present a hierarchy of the topics. Glynn and DiVesta (1977) found that when the students in their study carefully read outlines prior to reading text, overall recall was significantly increased. They reasoned (1977, p. 89) that an

> outline presented prior to reading the text alerts the learner to the inherent organisation of the content as well as the salient topics within the text. It primes the learner for the identification of relevant subsumers under which new material can be located for meaningful storage. . . .

In more recent terms, it can be argued that when students read outlines prefacing text, the outlines activate relevant schemata into which ideas in the to-be-read passage will be assimilated. As noted by Derry (1984), reading following the activation of relevant schemata should make the reading especially sensitive to incongruent or unusual ideas (see also Peeck *et al.*, 1982). In our view, if these incongruent ideas can be assimilated to the active schemata, they will be stored as "discrete schema corrections" (to use Derry's term). However, if the incongruent ideas cannot be assimilated to the relevant schemata, readers should identify them as false statements.

In other words, we argue that when students carefully read outlines prior to reading text, they will better be able to identify false statements in text than when no outline is read. While convincing incorrect statements may be added to a reader's knowledge base, obvious false statements will not be. In either event, however, such statements should be readily identifiable as false by readers. Our argument is based on the presumption that outlines activate relevant schemata prior to reading. These schemata, then, serve as the heuristic basis for assimilating or rejecting information contained in text.

A similar argument may be made for the role of headings in text. Headings are a form of text signal that have received only limited research attention (however, see work by Hartley and his colleagues, e.g., Hartley, 1980; Hartley *et al.*, 1980). Although headings can be generated in various ways (see Hartley and Burnhill, 1978), one common approach to the development of headings for text materials is to use the words and phrases generated in text outlines as the headings to be inserted in text. In our view, headings in text should, like outlines, activate relevant schemata in readers and guide their comprehension of text (Hartley *et al.*, 1980). From this perspective, students should be better able to identify false statements in texts containing headings than in texts without headings. Unlike outlines, however, headings probably provide only limited information to readers about the order of topics in a passage or their hierarchy because headings are each presented in isolation.

At this point, it is unclear to us whether headings alone or an outline alone would be more effective in aiding students' identification of false statements in text. Outlines provide information about the organisation of text not clearly available from headings. However, the role of knowledge about a text's organisation is not clear in terms of identifying false statements (although Baker and Wagner, 1987, have shown that false statements near the apex of a passage's hierarchy are more readily identified than false statements lower in the hierarchy of passage structure). Still, it may be that knowledge about text structure is important to the identification of false statements and that outlines will prove to be superior to headings as a means of facilitating students' identification of false statements. On the other hand, significant portions of outlines presented prior to reading may be forgotten as readers proceed through a text, where headings are distributed throughout the text and seem to be more likely to activate relevant schemata at each appropriate location in text. It may be that the distributed nature of headings would lead to greater facilitation of students' ability to identify false statements.

The current study focused on the effects of outlines and headings on readers' ability to identify false statements in text. Four conditions were employed: (1) a condition in which an outline prefaced the text, (2) a condition in which headings drawn directly from the outline were inserted in the text at appropriate locations, (3) a combined condition in which the outline prefaced the text and the headings were inserted, and (4) a control condition. In each instance, directions similar to those used by Baker (1985) were employed, asking readers to "underline anything hard to understand or incorrect". We predicted that students in the combined condition would find significantly more of the false statements than students in any of the other conditions. In addition, we predicted that similar levels of false statement identification would occur in the outline only and the headings only conditions. However, we also predicted that students in these conditions would identify significantly more false statements than students in the control condition.

## METHOD

### *Sample*

Sixty tenth and eleventh grade students enrolled at a laboratory school, who had completed a quarter-length course in modern European history the previous week, were selected. All had scored above 75 per cent on an examination covering the content used as reading materials in the current experiment. All students participated as part of the activities involved in a current social studies course. There were 34 females and 26 males. Ages ranged from 15 years 8 months to 17 years 1 month (M = 16 years 3 months).

### *Materials*

The text consisted of a 2,500-word summary of European events during the years 1946-1951 taken from Lane's (1985) recent volume. The text was re-typed on 11 × 8½ inch paper and duplicated for use. The text was also altered by eliminating all references to other materials in the book and by revising 15 sentences so that each of them presented clearly false statements. Examples of the original sentences and their false counterparts are presented below:

| | |
|---|---|
| Original Version: | The Berlin Airlift was precipitated by Soviet refusal to allow Allied traffic in or out of Berlin. |
| False Version: | The Berlin Airlift was precipitated by American and British refusal to allow Soviet traffic in or out of Berlin. |

| | |
|---|---|
| Original Version: | The Soviet army continued to occupy Poland at the outset of the 1950s. |
| False Version: | The British army continued to occupy Poland at the outset of the 1950s. |

In addition to the text, a 12-point outline employing three levels of organisation was prepared (see Appendix A). This outline prefaced the reading passage in some conditions. Finally, the text itself contained headings (taken directly from the outline) in some conditions but not others.

*Procedure*

All data were gathered in the students' regular classroom with conditions randomised within sessions.

Students were randomly assigned to conditions on the basis of the folder they received at the time of their experimental session. All students were asked to read the passage carefully and to underline any part of the passage that seemed hard to understand, contained inconsistencies, or was incorrect. Students in the control condition were given no other instructions. Students in the outline only condition were also asked to read carefully the outline prefacing the passage in their folders. To be sure that the outline was carefully attended, students in this condition were asked to copy the outline on a blank sheet of paper provided for that purpose. After reading and copying the outline, the students in this condition read the passage but were not allowed to look back at the outline.

Students in the headings condition were given instructions similar to students in the control condition but they were asked to read carefully the headings as they proceeded through the passage. To ensure that the headings were carefully attended, students in this condition were further asked to copy each heading in the margin near where it appeared in the passage. Students in the combined outline + headings condition received directions combining those given students in the outline only and headings only conditions.

All students were allowed to read at their own pace. When each finished reading, the folder was collected and the person was directed to complete a course-relevant reading. The next day, students were individually debriefed by the experimenter.

## RESULTS

Students' folders were examined to determine compliance with the directions. All students in the outline only and the outline + headings condition indeed copied the outline. In addition, all students in the headings only and the outline + headings condition copied the headings as directed. Then the folders were examined to determine which statements had been identified as false. Students were given credit for underlining any part of a sentence as having identified the entire sentence as false.

Only 17 truthful sentences were identified as false by students in the experiment. As these were equally distributed across the conditions (4, 4, 4, 5) and as no student incorrectly identified more than one sentence, these data were not further examined.

The number of false statements identified as false were tallied and entered into a one-way analysis of variance. The results, $F(3,59) = 45 \cdot 52$, $P < 0 \cdot 01$, $MSe = 4 \cdot 52$, indicated a significant difference among conditions. This analysis was then followed up via the Tukey HSD ($\alpha = 0 \cdot 05$) procedure. The results indicated that students in the combined condition identified significantly more false statements (M = 13·07, SD = 1·35) than students in any of the other conditions. No significant difference was observed between students in the outline only (M = 8·53, SD = 2·17) and the headings only (M = 8·47, SD = 2·57) conditions. However, students in each of these conditions identified significantly more of the false statements than students in the control condition (M = 3·93, SD = 1·92).

As a part of the individual debriefing sessions with students, the experimenter quizzed them on their knowledge of the false statements that had been inserted in the text. No discernible differences among the conditions in terms of knowledge of the content emerged (means for correctly responding to questions concerning the false statements ranged from 13·20 to 13·44).

## DISCUSSION

The results indicated that both headings and outlines significantly facilitated students' ability to detect false statements in text. Further, the combination of outlines and headings led to an extremely high rate of false statement detection.

The results of the control condition (approximately 26 per cent of the false statements were identified by the students) were highly similar to those reported for Baker's control subjects (e.g., 1985). The results of the outline only (57 per cent) and the headings only (56 per cent) conditions were comparable to the results reported by Baker when she informed students that false statements had been inserted in the text and that their task was to locate them. The results of the combined condition, however, indicated a rate of false statement detection (87 per cent) that was considerably greater than any previously reported in the literature.

Why should carefully reading outlines and headings have facilitated students' ability to identify false statements in text? Further, why should the combined procedure have had such a strong effect? An explanation congruent with accounting for why outlines and headings have an effect on readers' memory for text seems most straightforward. In this view, outlines and headings serve to activate relevant schemata. These activated schemata then serve as heuristics that guide the encoding of text information. Those segments of text that disagree with prior knowledge are discriminated from the remaining content and identified as false or at least incongruent statements.

It is possible that outlines have some of their effect as a result of providing readers with information about the structure of text while headings have their effect primarily via schema activation. This potential distinction cannot be addressed on the basis of the current study. However, the very powerful effect observed in the combined outline + headings condition might suggest a separate role for activated schemata and structural knowledge. That is, when students had to read both an outline prior to text and then headings inserted in the text their rate of identifying false statements was extremely high. This could have been due to a more powerful schema activation effect (activating the schemata prior to reading the text and then re-àctivating individual schemata at specific points of the text) but it also could have been due to the provision of structural knowledge about the reading passage in addition to schema activation. A more precise accounting will have to await additional research.

There seems little doubt that the procedures brought about the differences in students' rates of false statement identification. The random assignment of students to conditions assured equivalent knowledge of the content and the debriefings and actual course performance confirmed this. Consequently, it would seem that the applications of this research are straightforward. When teachers are concerned with students' ability to identify statements that disagree with their knowledge bases, they should devise outlines and/or headings for the reading materials and see that students carefully use these guides during reading.

Correspondence and requests for reprints should be addressed to Professor John A. Glover, Teachers College and Burris Laboratory School, Ball State University, Muncie, Indiana, USA 47306.

## REFERENCES

BAKER, L. (1984). Spontaneous versus instructed use of multiple standards for evaluating comprehension: effects of age, reading proficiency, and type of standard. *J. exp. child Psychol.*, 38, 289-311.

BAKER, L. (1985). Differences in the standards used by college students for evaluating their comprehension of expository prose. *Reading Res. Q.*, 20, 297-313.

BAKER, L., and ANDERSON, R. L. (1982). Effects of inconsistent information on text processing: evidence for comprehension monitoring. *Reading Res. Q.*, 17, 281-294.

BAKER, L., and BROWN, A. L. (1984). Metacognitive skills and reading. In PEARSON, P. D. (Ed.), *Handbook of Reading Research*. New York: Longman.

BAKER, L., and WAGNER, J. L. (1987). Evaluating information for truthfulness: the effects of logical subordination. *Memory Cogn.*, 15, 247-255.

DERRY, S. J. (1984). Effects of an organiser on memory for prose. *J. educ. Psychol.*, 76, 98-107.

EPSTEIN, W., GLENBERG, A. M., and BRADLEY, M. M. (1984). Coactivation and comprehension: contribution of text variables to the illusion of knowing. *Memory Cogn.*, 12, 355-360.

FOORMAN, B. R., and SIEGEL, A. W. (1985). *Acquisition of Reading Skills*. Hillsdale, N.J.: Erlbaum.

GLENBERG, A. M., WILKINSON, A. C., and EPSTEIN, W. (1982). The illusion of knowing: failure in the self-assessment of comprehension. *Memory Cogn.*, 10, 597-602.

GLYNN, S. M., and DIVESTA, F. J. (1977). Outline and hierarchical organisation as aids for study and retrieval. *J. educ. Psychol.*, 69, 89-95.

HARRIS, A. J., and SIPEY, E. R. (1983). *Readings on Reading Instruction*. (2nd ed.). New York: Longman.

HARTLEY, J. (1980). The effect of headings on children's recall from prose text. *Br. J. educ. Psychol.*, 50, 304-307.

HARTLEY, J., and BURNHILL, P. (1978). Fifty guidelines for improving instructional text. In HARTLEY, J., and DAVIES, L. K. (Eds.), *Contributions to an Educational Technology (Vol. 2)*. New York: Nichols.

HARTLEY, J., KENELY, J., OWEN, G., and TRUEMAN, M. (1980). The effects of headings on children's recall of prose. *Br. J. educ. Technol.*, 17, 164-174.

HUEY, E. B. (1908). *The Psychology and Pedagogy of Reading*. New York: Macmillan.

JUST, M. A., and CARPENTER, P. (1987). *The Psychology of Reading and Comprehension*. New York: Allyn and Bacon.

LANE, R. S. (1985). *Modern European History*. New York: Clelland.

PEECK, J., VAN DEN BOSCH, A. B., and KREUPELING, W. (1982). The effect of mobilising prior knowledge on learning from text. *J. educ. Psychol.*, 73, 771-777.

(*Manuscript received* 8*th December*, 1987)

## APPENDIX A

### OUTLINE EMPLOYED IN THE EXPERIMENT

- Post-war Europe
  - Reconstruction
    - The Marshall Plan
    - The Role of the Allies
  - Political Changes
    - Postwar Britain: Taking stock
    - Postwar France: Finding an Identity
    - A Divided Germany
  - NATO
    - A Response to Soviet Expansionism
    - An Iron Curtain
    - The Berlin Crisis

*Br. J. educ. Psychol.*, 58, 307-314, 1988

# REGIONAL DIFFERENCES IN INTELLIGENCE AND EDUCATIONAL LEVEL IN DENMARK

By T. W. TEASDALE
*(University of Copenhagen, Denmark)*

DAVID R. OWEN
*(Brooklyn College, City University, New York)*

AND T. I. A. SØRENSEN
*(Hyvidovre University Hospital, Copenhagen, Denmark)*

Summary. Intelligence test scores and educational levels of a representative sample of 71,768 draftees are reported for seven geographical regions in Denmark. Differences between the regions were found for both variables. However, analysis revealed that 92 per cent of the regional variation in intelligence test scores could be reproduced from the corresponding differences in educational level. These latter appear to be related to regional urbanisation and population density. The results are interpreted as suggesting that differences in intelligence between urban and rural areas arise from contemporary differences in the provision of education rather than, as is sometimes hypothesised, from the cumulative effect of previous generations of selective migration of relatively more intelligent people from rural to urban areas.

## INTRODUCTION

Variation in mean intelligence test scores for geographic regions has been reported for several countries. Lynn (1979) reviewed a number of British and Irish studies suggesting mean IQ variation among 13 regions of the British Isles with scores ranging from 96 for the Republic of Ireland to 102 for London and south-east England. The 1947 study of 11-year-old school children in Scotland (Scottish Council for Research in Education, 1949) showed a mean IQ of 105 for children resident in the four cities as compared to 101 for the counties. Lynn (1980), using data for male conscripts, also found regional differences in mean intelligence test scores across the 90 departments of France. In the United States, Powell and Steelman (1984) have reported on state-by-state variations in scores on the Scholastic Aptitude Test taken by high school students seeking admission to colleges and universities.

Kirkby (1980) has pointed out that the degree to which regional differences appear may depend greatly on how the regions are defined and that "there should preferably be some *a priori* reason for the taxonomy employed" (p. 335). Of the variables potentially underlying regional IQ variance, urban versus rural residence has perhaps been the most frequently examined.

In reviewing the early literature, Vernon (1979) notes that test scores have generally been lower in rural areas than in urban areas. He suggests two hypotheses to account for this phenomenon. The first is that it might be the result of selective migration of higher social class, better educated individuals from rural to urban areas, an effect which, furthermore, could cumulate over generations. This hypothesis is favoured by Lynn (1979, 1980) to account for his findings for the British Isles and France. It has likewise been proposed by Wolff (1979) to explain the 7- to 12-point IQ difference found in the United States between northern blacks and southern blacks, an explanation which as been contested (Kamin, 1980; Wolff, 1980). The greatest difficulty for the migration hypothesis is that adequate historical data concerning the educational and intellectual levels of the migrants of earlier generations are, at best, very difficult to obtain.

The second hypothesis is that regional differences in intelligence test scores might be the result of differences in contemporary social conditions, perhaps especially education, favouring urban areas. Here too, however, comprehensive data have been lacking hitherto, and, to our knowledge, no previous study has presented regional data on both individual test scores and educational levels.

Of the two hypotheses, it is the latter which we examine directly here. We have obtained data on regional differences in intelligence test scores and educational levels in Denmark. Support for the "social conditions" hypothesis obtains if the regional differences in intelligence are well predicted from the regional educational level differences.

## METHOD

The data for the present study were derived from the archives of the Danish draft board. Denmark has maintained military conscription uninterruptedly since 1946. The first stage of the induction procedure requires all males on attaining 18 years of age to appear before the draft board with a view to assessing suitability for military service. This appearance may be postponed for some years, especially where further education is being pursued, but occurs only very rarely after age 26. Individuals who are able to submit a certificate from a medical specialist documenting an illness which would disqualify them from military service are exempted from appearing before the board. Such illnesses include, for example, epilepsy, diabetes, and mental retardation of a severity requiring institutionalisation. Also exempted from appearing are those who volunteer for military service before the age of 18. Both forms of exemption, however, are rare, and about 95 per cent of all Danish males do appear in person before the draft board.

The draft board procedure includes the administration of a 45-minute group intelligence test, the Børge Prien Prøve (BPP; Rasch, 1980). This test was introduced in 1957 and continues to be used, unaltered, to the present day. It comprises four subtests, each of approximately 20 items: a letter-matrices test resembling Raven's Progressive Matrices, a verbal analogies test, a numerical sequences test, and a geometrical figures test. None of the subtests is multiple-choice, and the total score is the number of correct responses, that is, 0 to 78. We have elsewhere reported a correlation of 0·77 between this test and a Danish translation of the Wechsler Adult Intelligence Scale (Teasdale and Owen, 1986).

The procedure also incorporates recording the examinee's educational level. A nine-point scale is used ranging from 1 = minimum years of grade school to 8 = studentereksamen (equivalent to the British GCE advanced level) and 9 = university degree. More complete details of the scale are provided elsewhere (Teasdale and Owen, 1986), but it should be noted that for present purposes we have combined levels 8 and 9. Our information concerning the latter is rather incomplete since many university students elect to appear before the draft board prior to concluding, or even commencing, their degree courses (although they usually postpone the military service itself).

During the time period covered by the present study there were seven regional draft boards in Denmark. The boundaries shown in Figure 1 were those in effect from 1970 onwards. In that year some minor adjustments were made between the three regions constituting the Jylland peninsula. The figure also shows the location of major cities and towns and the population densities (per square kilometre) for each region. These data are derived from Danish census reports for 1968, which is the median year of draft board appearance for our sample (Danmarks Statistik, 1969).

FIGURE 1

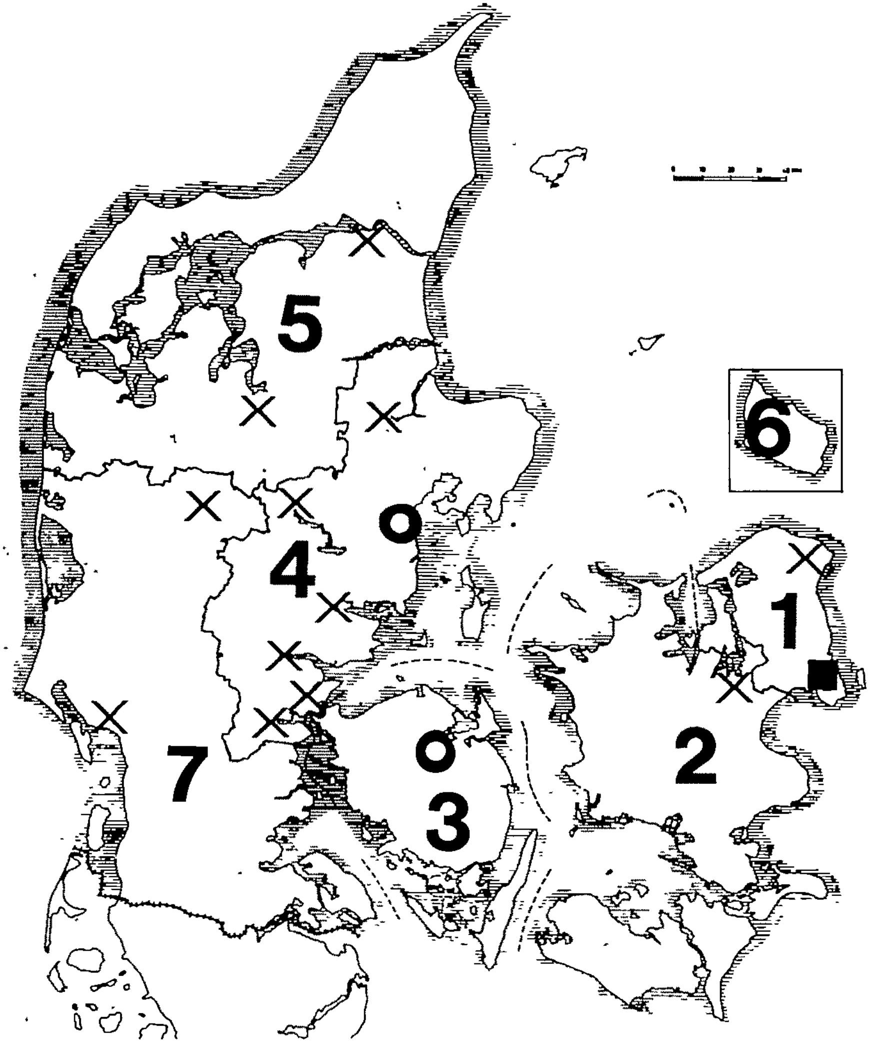
REGIONAL DIVISIONS OF DENMARK

| Region | Population Density (per km²) |
|---|---|
| 1. Northeast Sjaelland | 805·63 |
| 2. Remainder of Sjaelland and adjoining islands | 87·75 |
| 3. Fyn and adjoining islands | 123·08 |
| 4. Mid-Jylland | 107·32 |
| 5. North Jylland | 62·41 |
| 6. Bornholm | 81·61 |
| 7. South Jylland | 57·65 |

■ = city over 1,000,000 inhabitants (Copenhagen, region 1).

*O* = city over 100,000 inhabitants (Odense, region 3; Aarhus, region 4).

× = town over 25,000 inhabitants.



As part of a study also concerned with secular trends (Teasdale and Owen, 1987), we obtained draft board data on an approximately 10 per cent sample of all Danish males born between 1939 and 1958. This was accomplished by selecting the records for all males born on the first three days of each month throughout the 20-year period. From each record were noted the individual's intelligence test score, his educational level, and the region in which he had resided at the time of his appearance before the draft board. For ethical reasons, no individual identification was noted.

In all, data were collected on 76,111 individuals. Amongst these, 4,048 (5·3 per cent) had not appeared in person before the draft board, and test score and educational level were therefore lacking. A further 385 (0·6 per cent) records had incomplete data, leaving a total of 71,678 (94·2 per cent). The mean age for draft board appearance was 19·7 years (75 per cent were prior to age 20, and 99 per cent prior to age 26); 95 per cent of the appearances were between the years 1958 and 1977.

## RESULTS

Table 1 shows the means for birth year minus 1900 (including fractional year for birthdate; YR), educational level (EL) and test score (BPP), together with means and standard deviations for each of these three variables for all regions combined. Table 2 shows the corresponding univariate analyses of variance for each variable. There are significant differences among the seven regions on all three variables ($P < 10^{-5}$).

TABLE 1

EDUCATION LEVEL, COHORT, AND TEST SCORE MEANS BY REGION

| Region | N | EL | YR | BPP | Model BPP | Residual |
|---|---|---|---|---|---|---|
| 1 | 21,584 | 4·94 | 48·55 | 39·66 | 39·50 | 0·16 |
| 2 | 9,219 | 4·18 | 49·01 | 36·29 | 36·17 | 0·13 |
| 3 | 6,592 | 4·40 | 48·85 | 37·29 | 37·12 | 0·17 |
| 4 | 12,626 | 4·57 | 48·99 | 37·43 | 37·89 | −0·46 |
| 5 | 13,280 | 4·09 | 48·92 | 35·57 | 35·82 | −0·26 |
| 6 | 704 | 4·10 | 49·38 | 36·70 | 35·80 | 0·90 |
| 7 | 7,673 | 4·19 | 49·20 | 36·57 | 36·20 | 0·37 |
| Total | 71,678 | 4·48 | 48·86 | 37·50 | 37·50 | 0·00 |
| SD | | 2·08 | 5·57 | 12·16 | 8·72 | 8·48 |

*Notes:*

N = Number of individuals.
EL = Educational level.
YR = Birth year minus 1900.
BPP = Børge Prien Prøve raw score.
Model BPP = Predicted BPP derived from regression model using EL and YR.
Residual = BPP minus Model BPP.

The regional differences in birth year are very small ($\hat{\omega}^2 = 0{\cdot}002$), the difference between the minimum and the maximum being 0·83 years, equivalent to 0·15 standard deviations. The seven regional means are, however, negatively correlated with the regional population densities ($r = -0{\cdot}75$, $P < 0{\cdot}05$), suggesting either a relative deceleration of rural-to-urban migration and/or a relative acceleration of birth rate within the rural areas.

TABLE 2

EDUCATION LEVEL, COHORT, AND TEST SCORE ANOVAS BY REGION

| | Source | SS | df | MS | F |
|---|---|---|---|---|---|
| EL | Region | 8,447·47 | 6 | 1,407·91 | 333·42* |
| | Error | 302,636·39 | 71671 | 4·23 | |
| YR | Region | 3,633·65 | 6 | 605·61 | 19·56* |
| | Error | 2,219,298·41 | 71671 | 30·97 | |
| BPP | Region | 171,151·21 | 6 | 28,525·20 | 196·00* |
| | Error | 10,430,885·86 | 71671 | 145·54 | |
| Model BPP | Region | 157,597·03 | 6 | 26,266·17 | 355·46* |
| | Error | 5,295,987·97 | 71671 | 73·89 | |

*Notes:*
EL = Educational level
YR = Birth year minus 1900.
BPP = Børge Prien Prøve raw score.
* $P < 10^{-9}$

The regional differences in EL, while also small ($\hat{\omega}^2 = 0{\cdot}027$), are important, the difference between the minimum and the maximum being 0·85 points, equivalent to 0·41 standard deviations. There is a significant correlation between mean EL and mean population density across the seven regions (r = 0·86, P < 0·02).

The mean BPP differences between the regions are of about the same magnitude ($\hat{\omega}^2 = 0{\cdot}016$) as those for EL, the maximum-minimum difference being 4·78 or 0·39 standard deviations. Converted to a deviation IQ, this corresponds to a six-point difference between the highest scoring region and the lowest. The pattern of regional differences for the BPP test follows very closely that of EL, with a significant correlation (r = 0·91, P < 0·005) between regional BPP means and population densities.

We have earlier reported on secular trends in EL and BPP scores in the present data (Teasdale and Owen, 1987). We found that the BPP scores were well fit by a regression model containing linear and quadratic terms for EL plus an interaction term between EL and YR. A more elaborate computational algorithm, using double-precision arithmetic, has yielded very slightly improved regression weights, so for the following analyses the predicted Model BPP is now given by

$$\text{Model BPP} = 20{\cdot}85 + 3{\cdot}491 \times \text{EL} + 0{\cdot}1300 \times \text{EL}^2 - 0{\cdot}009814 \times \text{EL} \times \text{YR}$$

This model yielded a multiple correlation of 0·7172, and from it we concluded that the (upward) secular trend in BPP scores across the 20 years of our study could be well accounted for by the concurrent increases in EL.

For present purposes we have therefore employed this model to determine whether it can also account for the regional BPP differences. Table 1 includes the regional means for the predicted test score (Model BPP) and for residuals, that is, BPP minus Model BPP. Table 2 includes an analysis of variance of the predicted Model BPP values, and Table 3 shows an analysis of covariance (in which three degrees of freedom are taken up by estimates for the parameters in the regression equation for Model BPP). It can be seen that, even when covarying for the effects of EL, $EL^2$, and EL × YR, significant regional differences in BPP remain ($P < 10^{-9}$). These differences are very small in absolute size ($\hat{\omega}^2 = 0{\cdot}00053$), and Model BPP is accounting for a very large part of the between-region sample variance in BPP;

$$MS_{\text{MODEL BPP}}/MS_{\text{BPP}} = 26{,}266/28{,}525 = 92 \text{ per cent.}$$

The model residual means for the seven regions, shown in Table 1, do not correlate with the corresponding population densities ($r = -0{\cdot}04$, $P > 0{\cdot}92$).

TABLE 3

ANALYSIS OF COVARIANCE BY REGION

| | Source | SS | df | MS | F |
|---|---|---|---|---|---|
| Model BPP | 5,453,585·00 | 3 | 1,817,861·67 | 25,335·04* | |
| Region | 6,068·33 | 6 | 1,011·39 | 14·10* | |
| Error | 5,142,383·74 | 71668 | 71·75 | | |
| Total | 10,602,037·07 | 71677 | | | |

* $P < 10^{-9}$

## DISCUSSION

The present results show that there are, in Denmark, regional differences in IQ quite comparable with those which have been found in the British Isles (Lynn, 1979) and specifically within Scotland (Scottish Council for Research in Education, 1949). Earlier studies of regional IQ differences have, however, usually been unable to report on educational level differences, either because children were measured at the same age and therefore at the same ostensible educational level, or because, for adults, such information has not been available. In our data, not only do we find regional differences in educational level but also that such differences account for 92 per cent of the variance in regional IQ means. Furthermore, there is a very strong relationship between the average educational level of a region and its population density. From the analysis of covariance we can see that regional variation in intelligence test score, after conditioning on EL, becomes quite small; moreover, the remaining regional differences are not related with population density.

Taken together, these findings might be interpreted as follows. The provision of successive levels on the educational ladder becomes increasingly concentrated in urban rather than rural areas. This occurs for obvious economic and logistic reasons: establishments of further education tend to be placed where the size of the local and surrounding population warrants their presence. Thus, inhabitants of more densely populated regions have more ready access to further education. If prolonging formal education has a direct effect on IQ, then this would lead to higher mean IQs in densely populated regions. There is clear evidence that prolonged education does have a direct beneficial effect on IQ (Bouchard and Segal, 1985).

It would, however, be inappropriate to accept such an interpretation without some consideration of other factors. First of all, as noted above, some early studies have reported regional differences in IQ for school children at the same age who, therefore, do not differ in educational level as defined here, that is, ultimately attained level. This could implicate some qualitative differences in urban and rural educations. Such effects do not appear in our Danish data, perhaps because of a uniformity of educational quality across the country. One might speculate that regional qualitative differences are also disappearing elsewhere in western developed countries.

Another consideration is that of language. Vernon (1979) notes that in British studies the urban/rural comparison may be confounded with language differences, for example in Gaelic speaking or Welsh speaking areas. It is therefore worth pointing out that Danish is the sole native language in all parts of Denmark covered by the present study. (Residents of Greenland and the Faroe Islands are not subject to the Danish draft; subsequently, no comparable data exist for these regions.)

The main alternative to the interpretation that regional IQ differences are the result of regional educational differences is the possible role of migration. Lynn (1979, 1980) has shown that regional IQ means in both the British Isles and France correlate with estimates of population migrations since the 18th century. Since familial correlations in IQ are typically strong (Bouchard and McGue, 1981), then presumably a selective migration of more intellectually able people from rural to urban areas would lead to regional differences perpetuated through succeeding generations.

Some degree of migration has certainly occurred within our sample, but this has involved movement on the part of our subjects themselves, and it has often been in connection with the pursuit of the highest levels of education. During much of the time covered by the present study, the only Danish universities were in Copenhagen (region 1) and Aarhus (region 4); a university at Odense (region 3) opened in 1966. Since region is registered according to place of residence at the time of appearance before the draft board, there is a high probability that a university student would be registered in one of these regions irrespective of where he was born or grew up.

This does not necessarily mean, however, that selective rural-to-urban migration of the most intelligent has been compounding for centuries in Denmark and that it is differences in ancestor generations which have led to the present regional differences. First, the pursuit of higher education has not historically been the primary contributor to migration. The large part of the rural-to-urban migration which took place in Denmark in the last and present centuries has involved poorly educated rural workers moving to manual jobs in the industrial cities. Second, those individuals who move to the major cities to pursue higher education do not necessarily remain there after such education is completed. Third, if the effect had been cumulative over generations, then one would have expected to find regional mean intelligence test scores to be positively correlated with regional population densities even when controlling for educational level. Specifically, this should appear in Denmark since the densely populated areas, especially region 1 (Copenhagen), are those to which migration has historically occurred. We have, however, found no such correlation ($r = -0 \cdot 04$).

On balance, therefore, we are inclined to conclude that, in our data, regional IQ differences are indeed very largely the product of regional education level differences, themselves resulting from urbanisation factors. This conclusion is consistent with evidence we have presented elsewhere (Teasdale and Owen, 1984, 1987), indicating that education is one of the major environmental factors influencing IQ.

ACKNOWLEDGMENTS.—This work was supported by Grant 14-4092 from the Danish Social Science Research Council. We thank the Danish Ministry of Defence, Centre for Leadership and the Danish Ministry of the Interior, Department of Military Conscription, for permission to collect the data. The help and assistance of Departmental Leader Ketty Hansen were particularly appreciated. We are also indebted to Ms. Anna-Lise Petersen for clerical assistance.

Correspondence and requests for reprints should be addressed to T. W. Teasdale, Institut for Klinisk Psykologi, University of Copenhagen, Njalsgade 90, 2300 København S, Denmark.

## REFERENCES

BOUCHARD, T. J., and MCGUE, M. (1981). Familial studies of intelligence: a review. *Science*, 212, 1055-1059.

BOUCHARD, T. J., and SEGAL, N. L. (1985). Environment and IQ. In WOLMAN, B. J., (Ed.), *Handbook of Intelligence: Theories, Measurements, and Applications*. New York: Wiley.

DANMARKS STATISTIK (1969). *Statistisk Aarbog*. Copenhagen: Danmarks Statistik.

KAMIN, L. J. (1980). Selective migration — again. *Intelligence*, 4, 161-164.

KIRKBY, A. (1980). A critical comment on "The social ecology of intelligence in the British Isles." *Br. J. soc. clin. Psychol.*, 19, 333-336.

LYNN, R. (1979). The social ecology of intelligence in the British Isles. *Br. J. soc. clin. Psychol.*, 18, 1-12.

LYNN, R. (1980). The social ecology of intelligence in France. *Br. J. soc. clin. Psychol.*, 19, 325-331.

POWELL, B., and STEELMAN, L. C. (1984). Variations in state SAT performance: meaningful or misleading? *Harv. Educ. Rev.*, 54, 389-412.

RASCH, G. (1980). *Probabilistic Models for Some Intelligence and Attainment Tests* (revised). Chicago: University of Chicago Press.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1949). *The Trend of Scottish Intelligence: A Comparison of the 1947 and 1932 Studies of the Intelligence of Eleven-year-old Pupils*. London: University of London Press.

TEASDALE, T. W., and OWEN, D. R. (1984). Heredity and familial environment in intelligence and educational level — a sibling study. *Nature*, 309, 620-622.

TEASDALE, T. W., and OWEN, D. R. (1986). The influence of paternal social class on intelligence and educational level in male adoptees and non-adoptees. *Br. J. educ. Psychol.*, 56, 3-12.

TEASDALE, T. W., and OWEN, D. R. (1987). National secular trends in intelligence and education: a twenty-year cross-sectional study. *Nature*, 325, 119-121.

VERNON, P. E. (1979). *Intelligence: Heredity and Environment*. San Francisco: Freeman.

WOLFF, J. L. (1979). Selective migration among southern blacks: a reinterpretation of Lee (1951). *Intelligence*, 3, 139-148.

WOLFF, J. L. (1980). Selective migration — still a viable hypothesis. *Intelligence*, 4, 165-170.

(*Manuscript received* 28*th July*, 1987)

*Br. J. educ. Psychol.*, 58, 315-329, 1988

# ARE DISABLED READERS DELAYED OR DIFFERENT? AN APPROACH USING AN OBJECTIVE MISCUE ANALYSIS

By YVE WILLICH, MARGOT PRIOR, GEOFF CUMMING AND TOM SPANOS
*(Department of Psychology, La Trobe University, Victoria, Australia)*

Summary. The question of whether the reading behaviour of disabled readers is qualitatively different from that of normally reading children has as yet received no clear-cut answer. An objectively scored and simplified version of the Goodman miscue analysis system was developed as a means of comparing the use of various reading cues by Grade 4 disabled readers and Grade 2 normal readers whose reading level was psychometrically equivalent. The disabled readers were not only delayed but also qualitatively different from controls in that they made significantly less use of syntactic, semantic and text meaning cues. They also differed in self-correction behaviour with the disabled readers tolerating much more nonsense in their oral reading.

## INTRODUCTION

A variety of theoretical positions have guided research aimed at comparing normal disabled readers. Not surprisingly, the area is characterised by a very inconclusive pattern of results which are extremely difficult to synthesise; specific comparisons of actual reading behaviour in a qualitative sense have been relatively rare.

Goodman (1964, 1965, 1976) and others have promulgated a language-based approach to the investigation of the reading process emphasising the reader as a "language user" who makes use of a set of skills to organise and gain meaning from oral language and applies these same skills to reading.

Goodman (1964) has claimed that deviations from the text in oral reading, which he termed "miscues", are "windows to the reading process"; and that analysis of miscues would show the extent to which the reader was using each of the cues available in the text to produce the response. The Reading Miscue Inventory (RMI) was developed by Goodman and Burke (1972) as a test instrument for the analysis of oral reading miscues which would allow insight into the pattern of cue usage. Reading errors (i.e., miscues) are classified in terms of the rated influence of surface (phonic, graphic) and contextual factors in their production.

Clay (1968, 1969) also uses detailed analysis of errors. She found that at every grade level poor readers, at least in New Zealand schools, made less use of semantic and syntactic information than good readers and that they relied more on graphic and phonic cues. These differences increased with the age of the subjects (see also Murray and Maliphant, 1982). A similar pattern of results was reported by Goodman and Gollasch (1980) and by Goodman and Burke (1972), who noted additionally that non-word substitutions made by poor readers increased with age. Wixson's (1979) review of cue usage studies concluded that reliance on graphic and phonic cues decreases with age in proficient readers but not in poor readers. However, Juel (1980) reported that poor readers relied on both graphic and context cues and others (e.g., Perfetti *et al.*, 1979; Stanovich *et al.*, 1981) have claimed that poorer readers are *not* less able to use contextual or "top-down" processing. This lack of concordance emphasises the influence of age, task, situational effects and instructional methods on reading performance (Allington and Strange, 1977; Biemiller, 1979; Potter, 1982; Francis, 1984).

Despite good ecological validity and considerable research background, miscue analyses have not produced consistent results that permit generalisations concerning the developmental reading process or that provide prescriptions for better reading instruction. Leu (1981) has drawn attention to some of the theoretical and practical short-comings of miscue anlaysis including: inadequate definition of error categories, lack of consideration of the influence of text difficulty on resultant miscues, problems of multiple source errors, and the unassessed relationship between oral and silent reading, i.e., the issue of generalisability to "normal" reading practice. In addition, there is the critical question of the extrapolation that must be made from the analysis of reading errors to the understanding of accurate reading, an issue which is basic to the miscue logic. It is also unclear how consistent any child's strategies are from task to task or from hour to hour. There are further problems specific to Goodman's miscue analysis approach, including his value-laden assessment of errors and the over-complex error analysis categories and measures. The aim of the research reported here was to make a first step in enhancing the usefulness of miscue analysis, and to use the modified miscue procedure in an investigation of the nature of reading disability.

There are four aspects of this latter investigation. The first relates particularly to theories of reading disability, i.e., the question of whether disabled readers are qualitatively different from normal readers in their reading strategies, or whether they are simply delayed in acquisition of skills. Recently emerging developmental theories of reading (e.g., Marsh *et al.*, 1981; Frith, 1985) propose a series of stages from early logographic or whole word strategies through alphabetic or word analysis strategies to a final combined or orthographic stage, but say relatively little about the influence of context in this process. There is no consensus as to which approach best fits the progress of the normal reader. It is not surprising, therefore, that there is even more disagreement about disabled readers.

The "similar sequence" hypothesis of Weisz and Zigler (1979) which claims that intellectually disabled children progress through the same stages of cognitive development as normal children, albeit at a slower rate and with a lower ceiling, may also have heuristic value for the study of reading disability. The similar sequence model would predict that disabled readers matched on reading age level should read in the same way as normal reading controls. On the other hand, if their reading strategies are different or deviant they would appear dissimilar in the way they read. This question of delay versus difference has not been systematically examined and, although Leu's review (1981) suggests inconsistency in the evidence derived from miscue analyses especially relating to differences between good and poor readers, others (see Wixson, 1979) have synthesised findings which do permit some predictions for this study. Specifically, it was hypothesised that younger proficient readers would show a greater use of linguistic and contextual cues than disabled readers and that the latter would be more reliant on graphic and phonic (surface) cues than the proficient group, i.e., they would be different in addition to being delayed.

The other three issues are methodological ones, with the first being concerned with selection of appropriate control groups for research focused on the assessment of the characteristics of disabled readers. Disabled and normal groups have traditionally been matched on the basis of chronological age with IQ controlled. However, results from this design are confounded by the fact that the reading ages of such groups are obviously different and almost certainly reflect differential amounts of exposure to reading material and differential reading practice. The design also cannot speak directly to the delayed or different question that is of central interest here. Matching groups on reading age, however, does permit this question to be addressed.

A further methodological concern is with a problem apparent in previous miscue studies where the difficulty level of the reading material providing the raw material of analysis was not controlled. Degree of difficulty of reading material has varied across groups who are then compared on the basis of their error types. The different texts may facilitate or inhibit the use of different cue systems, leading to group differences which may be a function of the material rather than child reading strategies. This consideration is of particular importance in the light of Hood's (1982) findings that the degree of difficulty of the reading material can systematically affect cue usage and the resultant miscue analysis. Even when *degree* of difficulty for different reader ability groups is controlled, the problem of different texts for different groups persists. Thus, the more sophisticated text given to normal ability groups is likely to be more interesting, incorporate more complex sentences, contain more irregular words and hence facilitate the use of language based cues rather than graphic or phonic cues. Disabled readers, provided with much simpler text, shorter sentences, and a higher incidence of regular words, may be more prone to an approach with greater reliance on graphic and phonic information. An advantage of using reading age matched controls is that the same passages can be given to both experimental and control children.

The final issue relates to difficulties associated particularly with the Goodman RMI and the problem of error or miscue classification. In the study to be reported here, some revision of the Goodman RMI was undertaken to clarify the scoring procedure. The revision was guided by principles of parsimony and objectivity. Goodman's nine categories for miscue analysis, viz., dialect variation, graphic similarity, sound similarity, intonation shift, grammatical function change, miscue correction, grammatical acceptability, semantic acceptability and meaning change, were reduced to five categories — graphic and phonic similarity, syntactic, semantic and text meaning acceptability — on the grounds that there is considerable and controversial overlap among the nine original Goodman categories. Objectivity was sought by the introduction of a documented, quantitative scoring system which may be used in an unequivocal way by any reading researcher without the extensive training considered necessary for RMI usage and which still relies on subjectivity. Provision for the analysis of miscue correction was included but treated separately.

For this study, the disabled reader group was characterised by a delay in reading achievement of two years (by age and by grade level); average or above intelligence; similar home background and school learning experience to the normal reader control group; and absence of any known biological factor predisposing to reading disability. They might therefore be called dyslexic children although none had been formally diagnosed as such. They would also fit the Rutter and Yule (1975) criteria for specific reading retarded children. The aim of this study then was to compare the reading behaviour of such children with that of normal readers.

## METHOD

### *Sample*

The reading disabled group comprised 13 children from Grade 4 in a Melbourne suburban primary school: 11 boys and two girls with a mean chronological age (CA) of 9 years 6 months (SD 4 months). The normal reader group (11 boys and 2 girls) were from Grade 2 and had a mean CA of 7 years 5 months (SD 4 months). The two groups were equivalent on reading age: all 26 children had reading achievement scores which fell in the stanine range of 6-8 on level BB of the Australian Council for Educational Research, Primary Reading Survey, in both Word Knowledge and Comprehension Subtests. These scores are average or moderately above average for Grade 2 children according to State norms. All children came from the same middle class suburb, had been at the same school from first enrolment and had been exposed to the same method of initial reading instruction, consisting of a combination of a phonics programme with additional whole word recognition activities and the use of story books. All children had been screened for visual or hearing impairment by the school medical service and were in addition screened for auditory problems prior to the study with the ACER Auditory Discrimination and Initial Consonants Test. Non-verbal reasoning ability as assessed via the Raven's Coloured Matrices were average or above (range 50th to 95th percentile in both groups). None of the reading disabled group had received any systematic or consistent remedial treatment.

*Materials*

The Word Knowledge and Comprehension subtests of the ACER Primary Reading Survey, Level BB (suitable for children in 2nd grade) were used as the reading achievement measure. The Primary Reading Survey has been developed for Australian primary school children and assesses reading comprehension and word knowledge through grades 1 to 6. Word knowledge is assessed by having the child read a word and then, from three choices next to it, select the word closest in meaning to the target word. In the comprehension subtest, short passages of increasing length are presented, each with a number of multiple choice questions. Students are asked to choose the correct answer from three choices. Most questions ask the child to recognise a fact from the passage although one or two of the later questions ask for an inference based on the passage material. The passages are simple narratives or adaptations of sections of children's stories. Raw scores are converted to stanines and stanine scores provided the basis for matching of the two groups of readers in this study. It will be noted that no oral reading is required in this test. The Raven's Coloured Progressive Matrices (sets A, AB, and B) were used as the measure of general ability. Comprehension of oral language was assessed with the short form of the Token Test (De Renzi and Faglioni, 1978). Goodman and Burke's (1972) Reading Miscue Inventory (RMI) was used for instructing and testing subjects and for recording oral reading miscues.

All children read the fourth story ("War on Small Deer") from the RMI set of passages which have been constructed to take into account difficulty, word frequency and interest level. Selection of this RMI excerpt was made by giving several alternative passages to Grade 2 children who were not to be included in the control group and selecting that one which was of an optimum level of difficulty for that reading achievement level, in that it was easily understood while still generating miscues. Having established the fourth story as appropriate, class teachers were consulted to ensure that children had not previously encountered this story.

A modified version of the RMI scoring guide was developed for scoring miscues (see Appendix) with a numerical system to indicate the extent of cue use. The scoring for all cue categories (graphic, phonic, syntactic, semantic and text meaning) was identical and ranged from a score of one for minimal use of a cue to a score of three for maximum use. In contrast to the standard RMI procedure, detailed rules were stated for the allocation of scores to each miscue which obviated the need for scorers to rely on subjective judgment or level of familiarity with the procedure in order to score miscues. Table 2 provides examples. The "Strength and Weakness" profiles used in the RMI were not incorporated into the modified guide in order to avoid the possibility of theory laden judgments regarding the relative merits of different reading approaches. As an alternative measure of the relationship between strategy use and "self-correction", the modified guide provided a "failure to correct" score. This score represented the number of miscues which lacked syntactic, semantic or text meaning acceptability and were not corrected by the reader. This score then indicated the extent to which the reader tolerated material that violated the deeper language cues in his or her oral reading. Details of scoring procedures with examples are provided in the Appendix.

*Procedure*

Every reading retarded child from Grade 4 — as nominated by the teachers — and all Grade 2 children at the selected primary school were tested with the ACER Primary Reading Survey, Level BB. Scores on this test were used to achieve close pairwise matching of Grade 2 normal and Grade 4 disabled readers on reading achievement level (see Table 1). The two groups did not differ significantly on either word knowledge or comprehension scores. Children were also matched by sex. They were then given the auditory screening tests and the Token Test individually, and the Raven's Test was administered to each grade level separately as a group test. No child selected by reading score had to be eliminated as a result of screening tests. The mean Token Test error score was 0·8 for Grade 4 children and 1·0 for Grade 2 children.

TABLE 1

MATCHING DATA. RAW SCORES OF MATCHED PAIRS ON PRIMARY READING SURVEY (2ND GRADE LEVEL [BB])

| Grade Level | Pair No. | Sex | W.K.+ Score | Comp.++ Score |
|---|---|---|---|---|
| 4 | 1 | M | 16 | 15 |
| 2 | | | 18 | 14 |
| 4 | 2 | M | 15 | 15 |
| 2 | | | 16 | 17 |
| 4 | 3 | M | 18 | 18 |
| 2 | | | 17 | 19 |
| 4 | 4 | M | 17 | 11 |
| 2 | | | 17 | 11 |
| 4 | 5 | M | 16 | 11 |
| 2 | | | 16 | 11 |
| 4 | 6 | M | 19 | 9 |
| 2 | | | 17 | 8 |
| 4 | 7 | M | 18 | 19 |
| 2 | | | 18 | 19 |
| 4 | 8 | M | 14 | 17 |
| 2 | | | 15 | 18 |
| 4 | 9 | F | 18 | 19 |
| 2 | | | 18 | 19 |
| 4 | 10 | F | 14 | 12 |
| 2 | | | 12 | 14 |
| 4 | 11 | M | 13 | 17 |
| 2 | | | 14 | 19 |
| 4 | 12 | M | 16 | 20 |
| 2 | | | 16 | 19 |
| 4 | 13 | M | 19 | 13 |
| 2 | | | 18 | 15 |

+ Word Knowledge ++ Comprehension

All children then read the selected story aloud, after instructions as detailed in the RMI procedure. Recording of miscues was carried out as outlined in the RMI on the sheets provided, and in addition the oral reading was recorded for later replaying and checking of miscues.

### *Scoring*

All scoring was carried out by one experimenter and an independent assistant who had psycholinguistic training and who had no knowledge of the experimental design or of subject category. Inter-rater reliability for scoring was 0·94. The few disagreements were spread roughly evenly over the different cues and were resolved by discussion. In the light of this experience the scoring procedure was refined further, yielding the final form given in the Appendix. For each subject the first 25 miscues were compared with the text and a score of 1, 2, or 3 was allocated for each miscue on each of the five cue categories, to indicate the extent to which use of each miscue was evident in the reader response. A score of 1 indicated least use of a cue (i.e., larger difference between text and miscue), and of 3, most use of a cue (i.e., close correspondence between text and miscue).

Table 2 provides some examples of the way each miscue was scored for each of the five cue categories (See Appendix for full description of scoring guidelines). For the first miscue ("head"), the graphic and phonic similarity between the response and text is evident — and hence each category is scored 3, indicating maximum cue usage. To score syntactic and semantic acceptability, the entire sentence containing the miscue is considered. With this example the change from plural to singular, which violates the syntax of the sentence, caused this category to receive a score of 2; a score of 1, indicating minimal cue usage, is not appropriate as the syntactic violation is not total, i.e., a noun has still been used to replace a noun. The semantic acceptability, which by definition cannot exceed the syntactic acceptability, similarly receives a score of 2. In the final category of meaning retention, however, the score of 3 reflects the fact that when the test word is replaced by the response, the reader does not lose any of the information which is being conveyed in this sentence. The cue usage scores were totalled over miscues for each subject giving an overall score for each type of cue. Minimum possible cue use would be indicated by a score of 25 (25 miscues by score of 1) and maximum possible cue use would be 75 (25 × 3).

TABLE 2

EXAMPLE SCORING OF THREE MISCUES. FURTHER DETAILS OF THE SCORING PROCEDURE ARE GIVEN IN THE APPENDIX

| Reader's Response | Text | Graphic Similarity | Phonic Similarity | Syntactic Acceptability | Semantic Acceptability | Text Meaning Retention | Range of Cue Use | Correct | Fail to Correct |
|---|---|---|---|---|---|---|---|---|---|
| 1. head | heads | 3 | 3 | 2 | 2 | 3 | 3 | √ | |
| 2. became | began | 3 | 2 | 2 | 2 | 1 | 1 | | √ |
| 3. quirrl | squirrel | 3 | 3 | 1 | 1 | 1 | 2 | √ | |

In addition, for each miscue, a "fail to correct" score was recorded if the miscue scored below 3 on any one of the semantic, syntactic or text meaning acceptability and remained uncorrected. Total "fail to correct" scores were completely separate scores which indicated the extent to which what was being read was monitored according to context when the erroneous response was not changed. These were analysed separately.

## RESULTS

A Hotellings analysis which makes a multivariate comparison between two matched groups was carried out. It revealed a highly significant overall difference between the groups for cue use F (5, 8) = 10·52 P < 0·002. Subsequent univariate analyses showed significant differences between groups for the use of syntactic, semantic, and text meaning cues, but no significant differences for graphic or phonic cue use. F values and significance levels for these analyses are presented in Table 3.

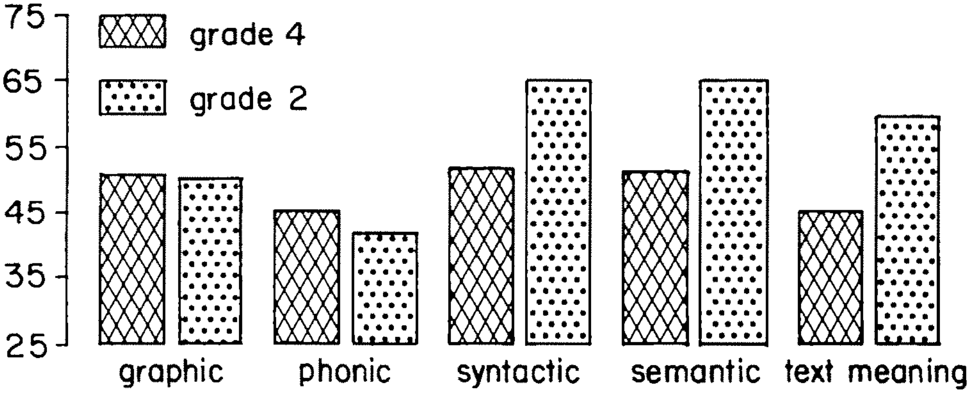
TABLE 3

UNIVARIATE F TEST RESULTS WITH (1, 12) DEGREES OF FREEDOM FOR CUE USAGE BY GRADE LEVEL OF SUBJECTS

| Cue/Variable | F Value | |
|---|---|---|
| Graphic | 0·17 | NS |
| Phonic | 0·38 | NS |
| Syntactic | 64·9 | P < 0·001 |
| Semantic | 51·1 | P < 0·001 |
| Text Meaning | 33·3 | P < 0·001 |

FIGURE 1

USE OF GRAPHIC, PHONIC, SEMANTIC, SYNTACTIC AND TEXT MEANING CUES BY DISABLED AND NORMAL READERS



The mean values for use of each of these cues by each grade level are presented in Figure 1, which shows that the use of syntactic, semantic and text meaning cues by the Grade 2 normal reading children was greater than that by the Grade 4 retarded readers.

Total scores for the usage of each cue by each subject pair were calculated. Groups were combined since correlations were similar for each group. These scores were used to obtain Pearson correlations amongst the five cues. The resultant coefficients are shown in Table 4.

It is evident that cue use scores for each of the three contextual cues (syntactic, semantic and text meaning) are highly correlated with each other, but with neither of the two surface cues (graphic and phonic). However, these surface cues are strongly correlated with each other.

TABLE 4

PEARSON CORRELATION CO-EFFICIENTS AND ONE-TAIL SIGNIFICANCE LEVELS FOR CUE USE BY GRADES 2 and 4 SUBJECT PAIRS

| Cue | Graphic | Phonic | Syntactic | Semantic | Text Meaning |
|---|---|---|---|---|---|
| Graphic | | 0·76<br>P = 0·001 | − 0·05<br>P = 0·43 | − 0·13<br>P = 0·35 | 0·14<br>P = 0·32 |
| Phonic | | | − 0·26<br>P = 0·23 | − 0·19<br>P = 0·26 | 0·01<br>P = 0·49 |
| Syntactic[1] | | | | 0·93<br>P = 0·001 | 0·55<br>P = 0·025 |
| Semantic[1] | | | | | 0·64<br>P = 0·009 |
| Text Meaning | | | | | |

*Footnote:*

The almost perfect correlation between scores in semantic *vs*, syntactic categories may suggest that one or other of these measures is redundant. Reference to the scoring system in the Appendix makes it clear that judgments of semantic acceptability are greatly influenced by syntactic constraints, as is the case, of course, in any psycholinguistic judgment of this type. Further modification of the RMI scoring system to allow for this would result in a more parsimonious system of analyses with little loss of information.

In order to obtain an indication of the range of cues used in generating miscues for each group, a mean range of cue use score was calculated for each subject. This range score was obtained by totalling for each miscue the number of cues (maximum 5) which scored 3, and averaging this number over the 25 miscues analysed. Hence, this score reflects the average number of cues that were strongly used in generating any miscue. A range score of 0 indicates strong use of no cue, while a range score of 5 indicates that all cues were strongly used and thus each assigned a score of 3. A matched pairs t-test which compared these range scores for the two groups revealed that the Grade 2 proficient readers had significantly greater range of cue use than the Grade 4 disabled readers (t = 3·07, df = 12, P < 0·005), (Grade 2 mean range of cue score — 2·35, SD, 0·33, Grade 4 mean = 1·80, SD 0·17).

A difference score (X) was calculated to indicate the extent to which contextual (or language) cue use exceeded surface cue use for each subject. This difference score was obtained by making the following calculation on the subject's five total cue use scores: X̄ = 1/3 (Syntactic + Semantic + Text Meaning Retention) − 1/2 (Graphic + Phonic).

With the exception of one score only, difference scores for Grade 2 proficient readers were all positive. A single sample test showed these Grade 2 scores to be significantly greater than zero (t = 6·43, df = 12, P < 0·005) and indicated that the use of contextual cues was significantly greater than the use of graphic and phonic cues for Grade 2 subjects (see Figure 1). A further test to compare the difference scores of Grade 2 proficient readers (X̄ = 16·73, SD = 9·18) with those of the Grade 4 disabled readers (X̄ = 7·17, SD = 5·43) showed the Grade 2 subjects to have significantly larger difference scores than the Grade 4 subjects (t = 4·3, df = 12, P < 0·005).

Although, of course, it is not independent of the preceding two tests, a further test was performed to compare the difference scores of the Grade 4 disabled readers with zero. (In this group, six of the 13 scores were negative.) The result showed no significant difference (t = 0·77, df = 12) indicating that there was no difference between the use of these two types of cue by this group.

A t-test for matched pairs was also used to compare the total "fail to correct" scores of the Grade 2 normal readers with those of the Grade 4 disabled readers (Grade 4: $\bar{X}$ = 13·77, SD = 3·6; Grade 2: $\bar{X}$ = 6·9, SD = 4·66). This analysis showed the difference between these two groups to be significant (t = 5·3, df = 12, $P < 0{\cdot}001$), with higher "fail to correct" scores in the disabled group.

Although all data used for the previous analyses were produced from the first 12 miscues made by each subject, the total number of miscues produced for the entire reading passage was also calculated for all subjects. A t-test was used to compare these totals for the two groups of subjects. The results showed no significant difference between the groups (t = 0·69, df = 12) in the total number of miscues produced (Grade 4: $\bar{X}$ = 37·38, SD = 8·1; Grade 2: $\bar{X}$ = 34·54, SD = 9·8).

## DISCUSSION

Although the small number of subjects included in this study limits the generality of the conclusions, the tight controls on subject selection, the paired subject design, and the control over the difficulty level of the source material permit some confidence in the findings in terms of the issues under investigation. The highly significant differences between the groups for the use of semantic, syntactic and text meaning cues suggest that disabled readers at this level are qualitatively different in their approach to reading, i.e., they are deviant in their strategies as well as delayed since they differ from younger but reading level equivalent children. This difference in the use of contextual information is consistent with the differences found in previous studies (Burke, 1976; Wixson, 1979). Our findings suggest that failure to make efficient use of contextual language information is truly characteristic of disabled readers, at least of this age level and instructional background, and not an artefact of the methodology employed in previous research.

However, in the pattern of graphic and phonic cue usage there is inconsistency between the results of this study and those employing chronological age matched designs. It is possible that the greater use of surface cues by disabled readers reported in previous studies (Wixson, 1979) was a result either of the differences in reading achievement level of the groups or of the differences in the reading materials used as tests for each group. No differences in the number of surface cues used between proficient and disabled groups was shown in the present study, thus disconfirming the claims of Christie (1981) and Goodman (1976) that proficient readers are attending to language cues *rather than* surface cues in predicting upcoming text; that is, increased attention to language cues was not accompanied by decreased attention to surface cues in our proficient reader group, at least by comparison with the disabled group. There was no evidence in this study that the use of one cue system substitutes for use of another, or that greater attention to context is accompanied by less attention to surface characteristics. It appears that, while equal use is made of the surface characteristics of print by our two groups of children, proficient readers gain *additional* information from the language based cue system in the text. Disabled readers are less efficient in their use of these deeper cue systems. Both group and pairwise comparisons showed this difference in cue usage and in addition the results indicated that the average number of cues used in generating a miscue was greater in the normal readers. The disabled readers then are more limited in their use of reading strategies than younger proficient readers, whilst the latter appear to make use of general knowledge of the reading process which results in greater use of all cue systems.

Such an interpretation is further supported by the evidence of differences in the self-correction behaviour of the two groups of children studied. This difference may indeed be the most important finding of the study since, while most of the data pertain to strategy use in predicting print, the self-correction results suggest deviance in the monitoring of responses which violated contextual information in disabled readers. The significantly greater correction of contextually inappropriate miscues by the normal readers indicates the use of such cues for monitoring as well as predicting what is being read.

It may be argued that monitoring is the more important of the two procedures in terms of gaining meaning from print since it determines the final representation of material to the reader. The finding here of tolerance of nonsense in what is being read, as evidenced by the "fail to correct" scores of the disabled readers, amplifies the findings of Goodman and Gollasch (1980) of greater frequency of non-word substitutions rather than omissions in such children.

It is perhaps surprising that investigation of cue use in *monitoring* text has been relatively neglected in reading research. Neither of the major Miscue Analysis Inventories (Goodman and Burke, 1972; Clay, 1979) includes a separate scoring procedure to analyse cue systems used to monitor reading. It might be suggested that the analysis of miscues which the reader fails to correct is a more direct and enlightening index of this factor than the traditional count of word substitution versus word omission (Goodman and Gollasch, 1980), especially since the frequency of word omissions seems highly dependent on teaching method and instructions at the time of testing. However it should be noted that Potter (1982) using a cloze task reported that self-corrections were not related to reading ability at least in 7-8 year-old children. Tendency to make self corrections may well be highly influenced by the teaching experienced by the child.

Much of the formal teaching of reading in the past has emphasised decoding by directing the attention of readers to the graphic and phonic cues in print. It is possible that proficient readers independently and/or intuitively abstract information regarding the linguistic framework for these aspects of reading from the more incidental, less formal aspects of teaching and other language experiences and as a consequence will develop a multiple strategy approach to reading (Francis, 1984). Those children who become disabled readers limit their attention to the cues to which they have been directed in formal reading instruction and, in the absence of training in the recognition and use of linguistic cue systems, fail to acquire the meta-knowledge which characterises their non-disabled peers. This may have been the case for the children participating in this study who had received initial reading instruction that emphasised phonics.

Alternatively, according to the theories supporting language delay as critical in the etiology of reading disability (e.g., Vellutino, 1979), the reading disabled child may begin to read without the necessary linguistic skills upon which to build word and print specific knowledge, i.e., he (usually he) is asked to run before he can walk and learns only to limp. If these early and basic differences are not remedied via reading instruction then this may result in the more restricted use of strategies seen in disabled readers in this study.

The conclusions of this study must be qualified by the need to rely on the basic assumption underlying all miscue analysis, that the nature of errors does give a valid indication of the nature of successful reading. In addition, whilst there is strong evidence of a qualitative difference in the reading behaviour of the disabled and normal readers in this study, results may not be generalisable to other age and reading level groups. Further studies should extend the comparison of reading strategies at various ages and levels of achievement and in children from different backgrounds, i.e., the developmental process of cue acquisition and use should be followed. From the results of this study it may be argued that an important component of reading remediation, at least at this age level, should be a language experience-based approach, with provision of opportunities for disabled and "at risk" readers to acquire the kinds of deeper knowledge of language which is evident in their non-disabled peers.

Acknowledgments.—The authors wish to acknowledge the contributions of Dr. B. McKenzie, Tess Gibbs and Carmel Allen, of the Department of Psychology, La Trobe University, to the development of this project.

Correspondence and requests for reprints should be addressed to Dr. Margot Prior, Department of Psychology, La Trobe University, Bundoora, 3083, Victoria, Australia.

## REFERENCES

Allington, R. L., and Strange, M. (1977). Effects of grapheme substitutions in connected text upon reading behaviours. *Visible Lang.*, 11, 285-297.

Biemiller, A. (1979). Changes in the use of graphic and contextual information as a function of passage difficulty and reading achievement level *J. read. Behav.*, 11, 307-318.

Burke, E. (1976). A developmental study of children's reading strategies. *Educ. Rev.*, 29, 30-46.

Christie, J. F. (1981). The effects of grade level and reading ability on children's miscue patterns. *J. educ. Res.*, 74, 419-423.

Clay, M. M. (1968). A syntactic analysis of reading errors. *J. verb. Learn. verb. Behav.*, 7, 434-438.

Clay, M. M. (1969). Reading errors and self-correction behaviour. *Br. J. educ. Psychol.*, 39, 47-56.

Clay, M. M. (1979). *The Early Detection of Reading Difficulties: A Diagnostic Survey with Recovery Procedures.* (2nd ed), Auckland, N.Z.: Heinemann Educational Books.

De Renzi, E , and Faglioni, P. (1978). Normative data and screening power of a shortened version of the token test *Cortex,* 14, 41-49.

Francis, H. (1984). Children's knowledge of orthography in learning to read. *Br. J. educ. Psychol.*, 54, 8-23.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In Patterson, K., Marshall, J., and Coltheart, M. (Eds.), *Surface Dyslexia.* London: Lawrence Erlbaum Associates.

Goodman, K. S. (1964). A linguistic study of cues and miscues in reading. Paper delivered at the American Educational Research Association, Chicago.

Goodman, K. S. (1965). A linguistic study of cues and miscues in reading. *Elementary English,* 42, 639-643.

Goodman, K. S. (1976). *Miscues: Windows on the Reading Process.* Detroit: University Press.

Goodman, Y., and Burke, C. L. (1972). *The Reading Miscue Inventory* New York: Macmillan.

Goodman, K. S., and Gollasch, E. V. (1980), Word omissions: deliberate and non-deliberate. *Read. Res. Q.*, 16, 6-30.

Hood, J. (1982). The relationship of selected text variables to miscue scores of second graders. *J. read. Behav.*, 14 (2), 140-158.

Juel, C. (1980). Comparison of word identification strategies with varying context, word type and reader skills. *Read. Res. Q.*, 15 (3), 358-376.

Leu, D. J. (1981). Oral reading error analysis: a critical review of research and application. *Read. Res. Q.*, 17 (3), 420-437.

Marsh, G., Friedman, M., Welch, V., and Desberg, P. (1981). A cognitive-developmental theory of reading acquisition. In MacKinnon, G., and Waller, T. (Eds.), *Reading Research: Advances in Theory and Practice, Vol. 3.* New York: Academic Press.

Murray, L. A., and Maliphant, R. (1982). Developmental aspects of the use of linguistic and graphemic information during reading. *Br. J. educ. Psychol.*, 52, 155-169.

Perfetti, C., Goldman, S., and Hogaboam, T. (1979). Reading skill and the identification of words in discourse extent. *Memory and Cognition,* 14, 273-282.

Potter, F. (1982). The use of the linguistic context: do good and poor readers use different strategies? *Br. J. educ. Psychol.*, 52, 16-23.

Rutter, M., and Yule, W. (1975). The concept of specific reading retardation. *J. Child Psychol. Psychiat. appl. Disciplines,* 16 (3), 181-197.

Stanovich, K., West, R., and Freeman, D. (1981). A longitudinal study of sentence context effect in second-grade children: tests of an interactive-compensatory model: *J. exp. Child Psychol.*, 32, 185-199.

Vellutino, F. R. (1979). *Dyslexia: Theory and Research.* Cambridge, Mass.: M. I. T. Press.

Weisz, J., and Zigler, E. (1979). Cognitive development in retarded and non-retarded persons: Piagetian tests of the similar sequence hypothesis. *Psychol. Bull.*, 86 (4), 831-851.

Wixson, K. L. (1979). Miscue analysis: a critical review. *J. read. Behav.*, 11, 163-175.

(*Manuscript received* 21*st August,* 1987)

## APPENDIX

*Scoring Guidelines*

A modification of Goodman and Burke's Reading Miscue Inventory is presented here. Each miscue is to be given a score of 1, 2 or 3 for each of the five cue categories:

1. Phonic Similarity.
2. Graphic Similarity.
3. Syntactic Acceptability.
4. Semantic Acceptability.
5. Retention of Text Meaning.

### GUIDE TO ALLOCATION OF SCORES

*Phonic Similarity*

A classification of errors according to similarity of sound between expected and observed response.

An indication of phonic similarity is the ratio between the number of phonemes in the observed response, which also occur in the expected response, and the total number of phonemes in the expected and observed response.

$$\text{i.e. P(Phonic)} = \frac{\text{S(Score)}}{\text{T(Total)}}$$

Where P = the index of phonic similarity

..... S = number of phonemes in the observed and expected responses which are the same.

..... T = total number of phonemes in expected and observed response.

The use of this formula will result in a range of indices from 0 to 1 for any miscue.

To translate these to the scoring range of 1 to 3 use the following table:

1. Score 1 for responses with an index from 0 to 0·33.
2. Score 2 for responses with an index from 0·33 to 0·66 inclusive.
3. Score 3 for responses with an index greater than 0·66.

*Examples*

| *Expected Response* | *Observed Response* | *Index* | *Score* |
|---|---|---|---|
| pet | dog | 0/6 = 0 | 1 |
| was | saw | 0/5 = 0 | 1 |
| *c*andle | *k*ite | 2/8 = 0·25 | 1 |
| l*a*ke | p*ai*nt | 2/7 = 0·29 | 1 |
| pu*ff* | laug*h* | 2/6 = 0·33 | 2 |
| *b*all | *b*at | 2/6 = 0·33 | 2 |
| b*a*b*y* | l*a*d*y* | 4/8 = 0·5 | 2 |
| *bl*ack | *ball* | 4/7 = 0·57 | 2 |
| be*d* | *said* | 4/6 = 0·67 | 3 |
| *come* | *cam*e | 4/6 = 0·67 | 3 |
| m*other* | br*other* | 6/9 = 0·67 | 3 |
| *skill* | *school* | 6/8 = 0·75 | 3 |
| *too*t | *two* | 4/5 = 0·8 | 3 |
| *two* | *too* | 4/4 = 1 | 3 |

(homophone may be distinguished by intonation)

*Special Cases:*

| | | | |
|---|---|---|---|
| 1. Substitution involving more than word | *b*rother | *b*ig boy | 2/10 = 0·2 (i.e., 1) |
| 2. Insertion of word | up stairs | up the stairs | Score 1 for "the" i.e., no similarity |
| 3. Omission of word | up the stairs | up stairs | Score 1 for "the" i.e., no similarity |

*Graphic Similarity*

Graphic Similarity is scored in the same manner as Phonic Similarity where all graphemes are counted and used as the denominator, and matched graphemes are used as the numerator.

*N.B.* Graphic Similarity scoring differs from Phonic Similarity in that while one point is given in the index for identical graphemes, 0·5 of a point is given for similar graphemes — i.e., same letter shape in same word position.

Hence, to score the response *boy* for text word *dog,* total graphemes = 6, therefore 6 becomes the index denominator; total number of graphemes the same — 2, therefore 2 is the first part of the numerator which indicates the graphemes which are the same. Total number of graphemes of same letter shape in same position = 4 × 0·5 = 2 (b/d and y/g).

This is added to the numerator indicating identical graphemes.

So the Graphic Similarity index $= \frac{2+2}{6} = \frac{4}{6} = 0{\cdot}66$

which corresponds to a graphic similarity score of 3 (based on index-score guide for phonic similarity). This score reflects the obvious high graphic similarity of the two words to a greater extent than would be gained by incorporating only identical graphemes in the numerator and hence provides a more sensitive indication of reliance on this aspect of print.

*Syntactic Acceptability*

Scoring 1-3

1 = No syntactic similarity or syntactically unacceptable response.
2 = Partial syntactic acceptability.
3 = Syntactically acceptable response.

Example test sentence: *The old man walked slowly.*

SCORE OF 1

1. *Omissions:* When the *remainder* of the sentence is not an acceptable structure. e.g. The old . . . walked slowly; . . . old man walked slowly.
2. *Single word substitutions:* When a nonsense word is used (except when this word retains *the same* tense ending, adjectival ending, adverb ending or plural ending as the word in the text) e.g., The old man wailly slowly.
   *or*
   When the word substituted violates the syntactic structure of the sentence, e.g., The old make walked slowly.
3. *Additions:* When the structure is changed to cause loss of acceptable syntactic structure. e.g., The old man boy walked slowly; The doll more doing to be winked slowly.

SCORE OF 2

1. *Omissions:* Where the omission causes the following word to be acceptable with the previous text *only.* e.g., The old man . . . slowly.
2. *Single word substitutions:* Where a real word substitution is acceptable with the previous part of the sentence only, or results in an incomplete sentence. e.g., The old man who slowly. The old man hasn't slowly. The old man was slowly.
   *or*
   Where verbs are read in the wrong tense or endings such as "ly", "ing" are omitted.
3. *Additions:* Where the added words are acceptable in structure with the previous part of the sentence only. e.g., The old man came walked slowly. The old man came over here walked slowly.

SCORE OF 3

1. Omissions: Where the syntactic structure of the remainder of the sentence is acceptable, e.g., The . . . man . . . walked.
2. Single word substitutions: Where a real word of the same part of speech is substituted to form a sentence which is syntactically acceptable, e.g., The old *mat* walked slowly.
   *or*
   Where a non-word with the *same* ending is substituted for the word in the text, e.g., The old man wacked slowly, or The old man walked sally.

3. *Additions:* Where the added words form a syntactically acceptable structure regardless of similarity to the text. e.g., The man walked over to the house slowly, or The mat was walking slowly, slowly by when the man saw it.

* Repetitions of the correct word are not scored as miscues, e.g., The man, man walked slowly by.
** Abbreviations not scored as miscues, e.g.: He wouldn't come *for* He would not come. Hippo for hippopotamus; T.V. for television. I'll for I will.

*Semantic Acceptability*
Scores between 1-3
1 = Not semantically acceptable.
2 = Partially semantically acceptable.
3 = Semantically acceptable.

*Example Sentence:* The little boy did not like school.

SCORE OF 1: *(Not semantically acceptable)*
1. *Omissions:* When the remainder of the sentence is *not* semantically acceptable. e.g., The little . . . did not like school.
2. *Single word substitutions:* When the substituted word violates the semantics of the sentence, i.e. a nonsense word is substituted. e.g., The little brog did not like school. The little boy did not like scheel.
3. *Additions:* When the added word is meaningless within the sentence. e.g., The little boys friend to me gave some school.; The little big did not like school.
   i.e., When the sentence is changed to cause a loss of acceptable semantic structure.

SCORE OF 2: *(Partial Acceptability)*
1. *Omissions:* Acceptable with what comes previously in the text but not in an entire sentence. e.g., The little boy did . . . school.
2. *Substitutions:* Where a real word substitution is acceptable with previous part of the sentence. e.g., The little town did not like school.
   *or*
   Where tense is changed, e.g., The little boy do not like school.
3. *Additions:* Where the added word(s) are acceptable with previous part of the sentence only. e.g., The little boy *was* did not like school.

SCORE OF 3: *(Semantically Acceptable)*
1. *Omissions:* Where the remainder of the sentence is acceptable. e.g., The . . . boy did . . . like school.
2. *Substitutions:* Where the sentence containing the substituted word is semantically acceptable as a whole. e.g., The little boy did.
3. *Additions:* Where the added word(s) form a semantically acceptable sentence regardless of the written text. e.g., The little boy did not like *going* to school.

*N.B.* As can be seen with syntactic scoring (above) semantic scoring can not be higher (i.e., a greater score) than syntactic scoring.

*Meaning Change*
Error to be considered within sentence. Scored 1-3.

*Score of 1*
Meaning conveyed by sentence containing the child's response *is different from* that conveyed by the text.

*Score of 2*
Meaning conveyed by the sentence containing the child's response (i.e. miscue) lacks substantial information contained in the text, or differs to some degree from the text, or cannot be determined.

*Score of 3*
Meaning of the sentence containing the miscue is the same as that of the text though minor qualifying information may be missing.

Examples: Text — *The little boy was very tired.*

*Score of 1:*
The little boy was very tried.
Then the boy wished very hard.
The little boy wasn't very tired.

*Score of 2:*
The little boy will be very tired.

*Score of 3:*
The boy was tired.
Then the boy was very tired.
The little boy were very tired.

*Self Correction*
This score is not one of the categories of reading cues. It reflects the extent to which the reader monitors miscues orally in his or her reading. It is scored by a YES or NO response only.

*Score YES:* in all cases where the reader abandons a miscue and gives the text response.

*Score NO:* where the reader abandons a miscue for an alternative which is not the text response, or, where the reader abandons a correct response for an alternative word, or, where a miscue remains unchanged.

If a word is repeated it is not scored at all as this does not constitute a "miscue" response.

*Failure to Correct*
This score is derived from the information given in the cue use scores for:

1. syntactic acceptability
2. semantic acceptability
3. text meaning retention

and the *NO* self correction scores.

For each miscue with a *no self correct* score — check to see if any of the three cue use scores listed above are below 3. If so, this miscue will be given an additional score of *failure to correct.*

If there is a *no* self correction score where syntactic, semantic and text meaning all have scores of 3 this miscue is not scored as failure to correct — as this score, in contrast to the simple self correction score, is an indicator of the relationship between language-based cue usage and correction behaviour.

*Br. J. educ. Psychol.*, 58, 330-337, 1988

# SLOW STARTERS AND LONG-TERM BACKWARD READERS: A REPLICATION AND EXTENSION

By ROB McGEE, SHEILA WILLIAMS AND PHIL A. SILVA*
*(University of Otago, New Zealand)*

SUMMARY. This study analysed data from a longitudinal investigation of reading ability in children from ages 7 to 13 years. Following Cox (1987), boys and girls with long-term backward reading (LTB) were compared with boys and girls who showed good reading progress at 13, despite being poor readers at 7 ("slow starters" or SS). These two groups were contrasted on various measures of cognitive development, behaviour and family background at age 7 in an effort to identify predictors of reading progress. Only maternal reading ability differentiated significantly between the SS and LTB groups.

## INTRODUCTION

WHILE reading ability tends to be relatively stable during early schooling (e.g., Williams and Silva, 1985), children do show differential rates of reading acquisition over time (Mackay *et al.* 1985). The question arises, therefore, as to whether it is possible to identify predictors of rate of reading progress. The identification of such predictors would be of practical value, particularly if they could be used to forecast which children are likely to suffer long-term reading disability.

Cox (1987) has recently compared two groups of reading disabled children first identified at age 7·5 years, and followed up at 15·5 years. One group, the so-called "Long-Term Backward" (LTB) readers, had an average reading age at 7·5 years of 5·5 years; eight years later, their average reading was only 8·1 years (all were below 9·3 years). The second group of so-called "Slow Starters" (SS) had an initial average reading age of 5·5 years; eight years later, their average reading age was 13·5 years (all were above 12·3 years). Although the Slow Starters had reading ages at age 15·5 which were sometimes below their chronological age, Cox argued that the group could be regarded as reading at an independent level. The Long Term Backward group, on the other hand, were markedly reading disabled at mid-adolescence.

Cox then compared the two groups on various predictors at age 6-7 years including vocabulary, non-verbal reasoning, behaviour at school, school attendance and home background. The results of a discriminant function analysis suggested that a combination of home background, vocabulary and problem behaviour achieved the best discrimination between the two groups of reading disabled children. However, univariate t-tests indicated significant differences between the groups only on the measure of home background and to a lesser extent, vocabulary. The statistical analysis was somewhat hampered by small sample size (N = 12 in each group) and Cox recommended replication of the study on larger samples.

This paper reports a replication and extension of Cox's (1987) research, with data from a longitudinal study of a large sample from Dunedin, New Zealand.

## METHOD

### *Sample*

The children were enrolled in the Dunedin Multidisciplinary Health and Development Study. This is a longitudinal investigation of their health, development and behaviour, and the study has been described by McGee and Silva (1982). To summarise, the sample were part of a cohort of children born between 1st April 1972 and 31st March 1973 at Queen Mary Hospital, Dunedin. The children were first followed-up at age 3 years, and 1,139 were eligible for inclusion in the study, in that they resided within the province of Otago. A total of 1,037 were assessed at age 3; the remaining 102 were either traced too late or their parents refused co-operation. The sample has subsequently been assessed at age 5 (N = 991), age 7 (N = 954), age 9 (N = 955), age 11 (N = 925) and age 13 years (N = 855).

---

* Rob McGee and Phil A. Silva are at the Department of Paediatrics and Child Health, Sheila Williams is at the Department of Social and Preventive Medicine.

The standard of living in New Zealand is high and the Dunedin sample is socio-economically advantaged when compared with the rest of New Zealand on an index of socio-economic status (SES: Elley and Irving, 1972). The sample is under-represented by Maori and other Polynesian children and is mainly European in origin.

### *Identification of LTB and SS readers*

Reading ability has been assessed at ages 7, 9, 11 and 13 years in the Dunedin study by the Burt Word Reading Test — 1974 Revision (Scottish Council for Research in Education, 1976). The test consists of 110 words graded in an approximate order of difficulty from easy (e.g., "to", "is", "big", "some" etc.) to hard (e.g. "autobiography", "microscopical", "phthisis", "subtlety"). The person tested is asked to read as many words as possible until attempting and failing 10 consecutive words. The score is the number of words pronounced correctly. A total of 397 boys and 393 girls had reading scores at all four ages. In addition, at age 7 years the children's reading ability was assessed by an 87-word prose reading test described by Silva (1981). The correlation between this test and the Burt test at age 7 was 0·86.

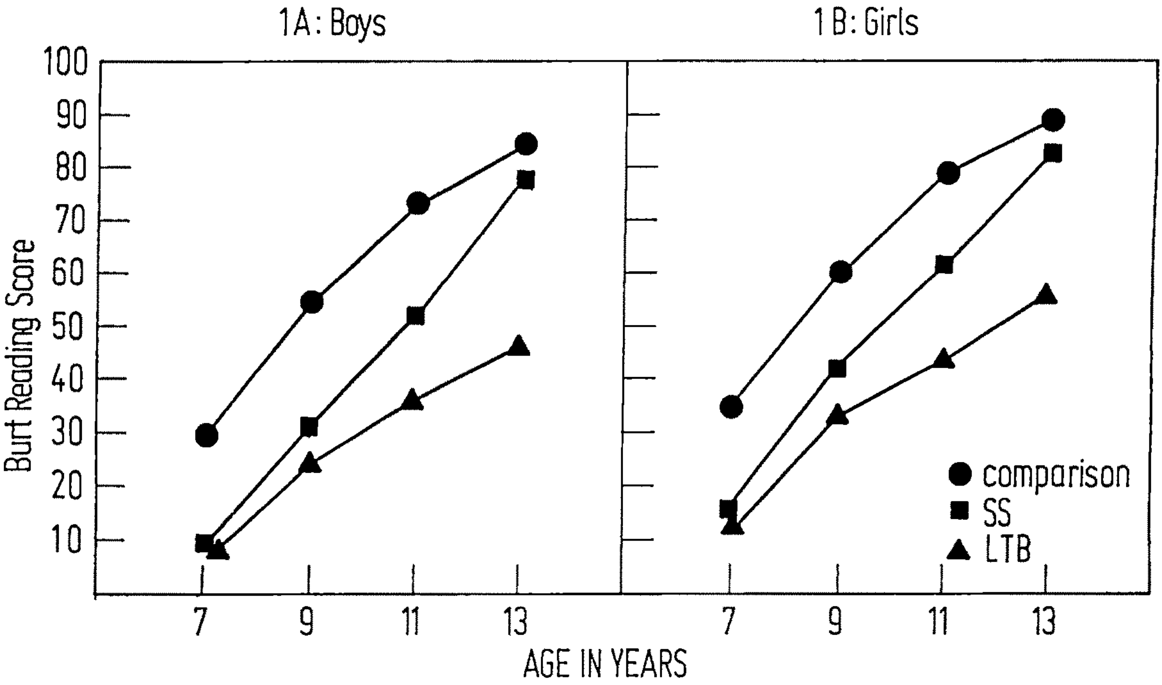
To identify poor readers at age 7, the criteria used was a reading score in the lower 15 per cent of the distribution for both the Burt test and the prose reading test. An additional criterion was having a WISC-R (Wechsler, 1974) IQ above 70 points. This procedure differs from that of Cox (1987) who used reading age norms to identify poor readers. Unfortunately, New Zealand norms for the Burt test could not be used in this way at age 7 (Gilmore *et al.*, 1981). However, by choosing poor readers on two different reading tests we believe we have identified groups of boys and girls who were reading significantly below the level of their peers. The 15 per cent criteria involved cut-off scores of 12 or less for the Burt test and 70 or less for the prose test in the case of the boys; the respective cut-off scores for the girls were 20 or less (Burt) and 77 or less (prose test).

The identification of slow starters and long-term backward readers was made with reference to the reading performance of these children at age 13 years, relative to their same-sexed peers. At age 13, the mean score on the Burt test of the 397 boys was 81·6 (SD = 18·8); the mean reading score of the 393 girls was 86·2 (SD = 15·4). The SS group were defined as those children who were poor readers at age 7 but by age 13 had a reading score *within* one SD unit of the mean for their sex (63 or more for boys; 71 or more for girls). This procedure identified 15 boys and 18 girls as slow starters. The LTB group were defined as those poor readers at age 7 who had a reading score at age 13 *more than* one SD unit below the mean for their sex. We identified 21 boys and 21 girls as long-term reading backward.

Figure 1 shows the means scores on the Burt reading test for the SS and LTB groups compared with the remaining children, from age 7 to 13 years. For the boys and girls in the LTB group, their reading ages at 13 were all below 11 years with an average between 8 years 3 months and 8 years 9 months (Gilmore *et al.*, 1981).

FIGURE 1

BURT READING SCORES FOR BOYS AND GIRLS IN THE SS, LTB AND REMAINDER GROUPS FROM AGES 7 TO 13 YEARS



*Measures*

The Dunedin study has collected a large amount of information concerning the characteristics of each child and his or her family. These include measures similar to those used by Cox (1987), although not identical. In order to examine the distinction between SS and LTB groups, the following variables were chosen. IQ was measured at age 7 by the WISC-R which allowed examination of Verbal (VIQ) and performance (PIQ) components as well as individual subscales. (The subscales of comprehension and picture arrangement were omitted because of time constraints.) Verbal comprehension (auditory reception) and expression were measured at age 7 using the relevant subscales of the Illinois Test of Psycholinguistic Abilities or ITPA (Kirk *et al.*, 1968). Speech articulation was measured using the Dunedin Articulation Check (Justin *et al.*, 1983), a 20-item measure scored for the correct articulation of 6 phonemes and 14 groups of sounds.

Behaviour problems at age 7 were assessed using the Rutter Child Scales A and B for completion by parent and teacher, respectively (Rutter *et al.*, 1970). These scales contain 31 and 26 items and list a variety of problems rated as "does not apply" — scored 0; "applies somewhat" — scored 1; and "certainly applies" — scored 2. A total problem score can be obtained by summing over items. School absences at age 7, obtained from teacher report, were also examined. School absence was expressed as number of days absent as a proportion of total number of possible days' attendance.

A measure of "family adversity" was calculated at age 7, based upon low SES (semi-skilled/unskilled father); large family size of 4 or more children; parental separation(s); low maternal mental ability; poor maternal mental health; and poor family relationships. McGee *et al.* (1985) fully describe this index. Each child was given a score from 0 to 6 on this index according to the presence (1) or absence (0) of the item. In addition to this general index of family adversity, the Family Environment Scale (FES: Moos, 1974) subscales of family "intellectual-cultural" orientation and "achievement" orientation were also specifically examined. These subscales provide a measure of the family interest in political, intellectual and cultural activities and the degree to which achievement is stressed. Also examined was the "family relations index" or FRI, a measure of the degree of perceived family support based upon the FES subscales of "cohesion", "conflict" and "expressiveness of feeling". The FES itself was usually completed by the child's mother.

Finally, a measure we believed to be of some interest was maternal reading ability, also assessed by the Burt test when the children were age 3. In assessing the women, the test commenced well into the list so that they did not have to read the full list of words. Nearly all women accepted being tested for reading in this way. However, those women who showed reading problems in the early part of the interview-assessment, or whose first language was not English, were not tested to avoid embarrassment for them. Reading scores were available for 98 per cent of the sample.

### *Procedure*

The general procedure at each phase of the assessments was the same. The children attended the Dunedin Unit within one month of their birthdays, and were assessed on a variety of health, cognitive, motor and behavioural measures. The assessments of IQ, verbal abilities, speech and reading were carried out by psychometrists trained in the use of these tests. During the early phases of the study, a parent, for the most part the mother, accompanied the child to the Unit. The mothers were also interviewed and detailed background information obtained on the child and his or her family. At the age 3 assessment, the mothers were administered the Burt reading test by the psychometrist.

## RESULTS

For the most part, the results were analysed by a series of 2 (Sex) by 3 (Group) analyses of variance (ANOVA), followed by post-hoc Scheffé contrasts where overall differences among the groups were demonstrated. Because of the relatively large number of significance tests, the Bonferroni procedure was used to control for potential inflation of the Type I error rate (Grove and Andreasen, 1982). Table 1 shows the mean scores for the measures of IQ, verbal comprehension and expression, speech articulation, behaviour problems, and the FES measures. As there were no significant Sex × Group interactions, the results are shown for boys and girls combined.

On all measures except those relating to the family environment, there were significant overall differences among the groups on ANOVA. However, comparison of the LTB and SS groups by Scheffé contrasts on the developmental and behavioural measures indicated no significant differences between these two groups, although the LTB group had lower scores on some of these measures. While this procedure could be perceived as relatively conservative, we believe it to be appropriate. Decisions to discriminate between slow starters and long-term backward readers need to be based on clear differentiations between these groups. This would be particularly so if differential intervention was to be considered. Both reading groups differed from the comparison group on these measures ($P < 0 \cdot 05$), indicating a generally lower ability associated with reading disability.

TABLE 1

MEAN SCORES FOR DEVELOPMENTAL MEASURES, BEHAVIOUR PROBLEMS AND FAMILY ENVIRONMENT IN THE SS, LTB AND COMPARISON GROUPS

| | | Group | | | |
|---|---|---|---|---|---|
| Measure | LTB (N=42) | SS (N=33) | Comparison (N=709[b]) | $\sqrt{MSe}$[a] | F (2,780) |
| VIQ | 91·8 | 96·4 | 107·8 | 14·0 | 34·23* |
| PIQ | 96·2 | 99·6 | 108·6 | 13·6 | 22·13* |
| Comprehension | 23·0 | 26·5 | 29·8 | 7·6 | 17·71* |
| Expression | 25·5 | 27·5 | 30·2 | 8·2 | 7·76* |
| Articulation | 11·3 | 12·2 | 15·0 | 3·3 | 33·69* |
| Behaviour: | | | | | |
| Parent Scale A | 11·1 | 12·9 | 8·4 | 5·3 | 15·14* |
| Teacher Scale B | 6·4 | 5·3 | 3·7 | 4·7 | 8·38* |
| School Absence | 6% | 5% | 6% | 5% | <1·00 NS |
| FES: | | | | | |
| Intellectual | 5·2 | 5·8 | 6·0 | 2·0 | 3·36 NS |
| Achievement | 4·7 | 4·6 | 4·2 | 1·8 | 2·68 NS |
| FRI | 18·7 | 18·8 | 19·4 | 3·5 | 1·11 NS |
| Maternal Reading | 93·8 | 99·4 | 102·1 | 8·5 | 19·34* |

* $P<0\cdot05$.

[a] MSe represents square root of the mean square error term for ANOVA.

[b] These numbers may vary slightly owing to some missing values.

Mackay *et al.* (1985) examined various subscore patterns on the WISC to see if they could distinguish between slow starters and a long-term failing group of readers. We also investigated performance on the WISC-R subscales of information, similarities, arithmetic, vocabulary, picture completion, block design, object assembly and coding. All subscales showed overall significant differences among the three groups with no significant Sex × Group interactions. However, there were no significant differences between subscale means for the SS and LTB groups (Scheffé test).

While there was no overall difference between the SS and LTB groups in terms of total behaviour problem score, it was decided to examine a measure of "hyperactivity" in particular. Other research from the Dunedin study (e.g., McGee *et al.*, 1985, 1986) has highlighted the relationship between reading disability and this type of problem behaviour. The measure of "hyperactivity" consisted of the sum of the individual items for "poor concentration", "very restless . . . hardly ever still" and "squirmy, fidgety child". A score of 4 or more (of 6) on the teacher scale was used to define a high score. Table 2 shows the numbers of children in each of the three groups at age 7 who scored 4 or more on teacher rated hyperactivity. These results were initially analysed using log-linear modelling procedures (Fienberg, 1978) testing for a significant Sex × Group interaction. This was not significant, $P > 0\cdot05$. An overall $\chi^2$ across the three groups was significant, however, with $\chi^2 = 23\cdot05$, 2 df, $P < 0\cdot05$. Post-hoc contrasts of each reading disabled group with the comparison group indicated that only the LTB group differed from the comparison group (Everitt, 1977). This suggests a higher rate of teacher rated hyperactive behaviours in the LTB group.

TABLE 2

PROPORTIONS OF BOYS AND GIRLS SCORING 4 OR MORE ON TEACHER RATED HYPERACTIVITY AND 2 OR MORE ON THE FAMILY ADVERSITY INDEX, AT AGE 7

| | LTB | SS | Comparison |
|---|---|---|---|
| A. Teacher Rated Hyperactivity (score 4+) | | | |
| Boys | 8/21 | 1/15 | 28/356 |
| Girls | 3/20 | 2/18 | 17/345 |
| B. Family Adversity (score 2+) | | | |
| Boys | 11/21 | 7/15 | 70/359 |
| Girls | 5/21 | 4/18 | 64/350 |

For the measure of family adversity, a score of 2 or more was considered as indicating some degree of adversity in the family environment. This applied to about 20 per cent of the sample. Table 2 also shows the numbers of boys and girls in the each of the three groups scoring 2 or more on the measure of family adversity. Once again there was no significant Sex × Group interaction, but an overall $\chi^2$ test across the three groups was significant with $\chi^2 = 12 \cdot 41$, 2 df, $P < 0 \cdot 05$. Post-hoc contrasts indicated that only the LTB group differed significantly from the comparison group. However, a direct comparison of the LTB and SS groups suggests that these two groups would not differ significantly in terms of family adversity (16/42 in the LTB group or 38 per cent *vs.* 11/33 in the SS group or 33 per cent). Both groups appear to have a higher rate of family adversity than the comparison group, with a rate of 19 per cent.

Finally, the measure of maternal reading ability indicated a significant overall difference among the groups and post-hoc comparison of the SS and LTB groups (99·4 *vs.* 93·8) was significant with $F = 7 \cdot 72$, $P < 0 \cdot 05$. That is, the mothers of those children in the SS group had significantly better reading ability than the mothers of children in the LTB group. In fact, the reading score of mothers in the SS group did not differ significantly from the remaining mothers in the comparison group. (There was one missing value for the mothers' reading score in the SS group and two missing values in the LTB group. Consequently, the respective sample sizes were 32 and 40 for this comparison. The standard deviations were 13·1 in the SS group, 15·1 in the LTB group and 9·6 for the remainder. While this suggests slightly more variation in scores in the two reading disabled groups, this may be a function of a ceiling effect in the remainder of the sample.)

Because of the potential significance of this finding, the distribution of maternal reading scores in the SS and LTB groups were examined. These distributions are shown in Table 3. While there were relatively few extreme low scores in either group, there were proportionally more women scoring between 71 and 100 in the LTB group compared with the SS group (48 per cent *vs.* 22 per cent). That is, the lower mean reading score in the LTB group does not appear to be the result of a few very low scores.

TABLE 3

DISTRIBUTION OF MATERNAL READING SCORES IN THE SS AND LTB GROUPS AT AGE 3

| | Reading Score | | | | |
|---|---|---|---|---|---|
| | 70 or less | 71-80 | 81-90 | 91-100 | 101-110 |
| SS | 6% | 0% | 6% | 16% | 72% |
| LTB | 7% | 5% | 10% | 33% | 45% |

## DISCUSSION

Overall, the findings of the present study indicate that measures of the child's cognitive ability, language abilities and speech, measures of behaviour at home and at school, and family environment and adversity do not discriminate in predictive fashion between children showing long-term reading disability and those who are "slow to start" reading. With respect to IQ and pattern of subscale scores on the WISC-R, the results confirm those of Mackay *et al.* (1985) who found that subscale groupings did not discriminate between SS and LTB groups.

To some extent such conclusions differ from those of Cox (1987) who claimed that measures of family background and child variables could be used to discriminate between SS and LTB groups. However, Cox used discriminant function analysis, a technique which is likely to be sample specific. The test of Cox's claim would be whether such discrimination can be replicated. Certainly, our results using univariate analyses do not support Cox's findings.

On the other hand, Cox's univariate analyses found the strongest differentiation between the SS and LTB groups on the measure of family background. This measure included ratings of availability of books, cultural experiences and level of parental interest in the child's education. It would appear that our results and those of Cox converge in this area. The strongest discriminating variable in our study between the SS and LTB groups was maternal reading level. While both groups seemed to have about equivalent levels of family adversity, maternal reading was poorer in the LTB group.

The Burt reading test was designed primarily as a measure of children's word recognition skills, so it could be argued that it is not an appropriate measure of adult literacy. However, although primarily designed for children and adolescents, the test does contain many words which are not in frequent usage. For example, of the 30 words at the "difficult" end of the test, 27 occur less than 15 per million in the Kucera and Francis (1970) word count. This would suggest that many of the words are not "childish" or necessarily easy for adults to pronounce. While this difference between the SS and LTB groups in mothers' reading is of some interest, the possible mechanism by which maternal reading level could influence the child's reading ability cannot be identified in our study. Nevertheless, other researchers have also found evidence for the importance of the child's home "literary" environment in achieving literacy (e.g., Hewison and Tizard, 1980; Share *et al.*, 1983). Taken together, this body of results may indicate that variables relating to literacy in the home are the best predictors as to whether a child showing early reading disability will achieve normal levels of reading skills in later life.

Finally, we agree with Cox that school-based interventions need to be supported by programmes to assist parents in taking an active role in the development of their child's reading skills. The results of this study suggest that the parents' own level of literacy is likely to play an important part in the development of such skills. A programme similar to the "cross-age" tutoring programme of Limbrick *et al.* (1985) may provide a technique by which both parent and child literacy are enhanced.

ACKNOWLEDGMENTS.—The Dunedin Multidisciplinary Health and Development Research Unit is supported by the Medical Research Council of New Zealand and involves several departments of the University of Otago. Much of the data are gathered by voluntary workers from the Dunedin community. The authors are indebted to the many people whose contributions make this ongoing study possible.

Correspondence and requests for reprints should be addressed to Dr. Rob McGee, Dunedin Multidisciplinary Health and Development Research Unit, Department of

Paediatrics and Child Health, Medical School, University of Otago, Box 913, Dunedin, New Zealand.

## REFERENCES

COX, T. (1987). Slow starters versus long-term backward readers. *Br. J. educ. Psychol.*, 57, 73-86.

ELLEY, W. B., and IRVING, J. C. (1972). A socio-economic index for New Zealand based on levels of education and income from the 1966 Census. *N.Z. J. educ. Studies*, 7, 155-167.

EVERITT, B. S. (1977). *The Analysis of Contingency Tables.* London: Chapman and Hall.

FIENBERG, S. E. (1978). *The Analysis of Cross-Classified Categorical Data.* Cambridge, Mass: MIT Press.

GILMORE, A., CROFT, C., and REID, N. (1981). *Burt Word Reading Test — New Zealand Revision. Teacher's Manual.* Wellington: New Zealand Council for Educational Research.

GROVE, W. M., and ANDREASEN, N. C. (1982). Simultaneous tests of many hypotheses in exploratory research. *J. nerv. ment. Dis.*, 170, 3-8.

HEWISON, J., and TIZARD, J. (1980). Parental involvement and reading attainment. *Br. J. educ. Psychol.*, 50, 209-215.

JUSTIN, C., LAWN, L., and SILVA, P. A. (1983). *The Dunedin Articulation Test.* Dunedin: Otago Speech Therapy Association.

KIRK, S. A., MCCARTHY, J. M., and KIRK, W. D. (1968). *The Illinois Test of Psycholinguistic Abilities.* Urbana, Ill.: University Press.

KUCERA, H., and FRANCIS, W. N. (1970). *Computational Analysis of Present-Day American English.* Providence, R.I.: Brown University Press.

LIMBRICK, E., MCNAUGHTON, S., and GLYNN, T. (1985). Reading gains for underachieving tutors and tutees in a cross-age tutoring programme. *J. child Psychol. Psychiat.*, 26, 939-953.

MACKAY, M. F., NEALE, M. D., and THOMPSON, G. B. (1985). The predictive validity of Bannatyne's WISC categories for later reading attainment. *Br. J. educ. Psychol.*, 55, 280-287.

MCGEE, R., and SILVA, P. A. (1982). *A Thousand New Zealand Children: Their Health and Development from Birth to Seven.* Special Report Series Number 8. Auckland: Medical Research Council of New Zealand.

MCGEE, R., WILLIAMS, S., BRADSHAW, J., CHAPEL, J. L., ROBINS, A., and SILVA, P. A. (1985). The Rutter scale for completion by teachers: factor structure and relationships with cognitive abilities and family adversity for a sample of New Zealand children. *J. child Psychol. Psychiat.*, 26, 727-739.

MCGEE, R., WILLIAMS, S., SHARE, D. L., ANDERSON, J., and SILVA, P. A. (1986). The relationship between specific reading retardation, general reading backwardness and behavioural problems in a large sample of Dunedin boys. *J. child Psychol. Psychiat.*, 27, 597-610.

MOOS, R., (1974). *Combined Preliminary Manual for the Family, Work and Group Environmental Scales.* Palo Alto., Calif.: Consulting Psychologists Press.

RUTTER, M., TIZARD, J., and WHITMORE, K. (1970). *Education, Health and Behaviour.* London: Longmans.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1976). *The Burt Word Reading Test* (1974 Revision). London: Hodder and Stoughton.

SHARE, D. L., JORM, A. F., MACLEAN, R., MATTHEWS, R., and WATERMAN, B. (1983). Early reading achievement, oral language ability and a child's home background. *Aust. Psychol.*, 18, 75-87.

SILVA, P. A. (1981). A simple prose reading test to identify seven year olds with reading problems. *Nat. Educ.*, 9, 36-38.

WECHSLER, D. (1974). *The Wechsler Intelligence Scale for Children — Revised.* New York: Psychological Corporation.

WILLIAMS, S. M., and SILVA, P. A. (1985). Some factors associated with reading ability: a longitudinal study. *Educ. Res.*, 27, 159-168.

(*Manuscript received 4th January,* 1988)

*Br. J. educ. Psychol.*, 58, 338-344, 1988

# YULE'S "HUMP" REVISITED

BY D. ALAN DOBBINS
*(University College of Swansea, Wales)*

SUMMARY. Reading and non-verbal intelligence tests were given to over 5,000 nine- to 10-year-old pupils to investigate whether there was evidence to support the existence of Yule's "hump" — a group of children with severe and specific retardation which is not just the lower end of a normal continuum (Yule *et al.*, 1974). The results are inconclusive but may, depending on interpretation, give some support to the notion of a small "hump".

## INTRODUCTION

The notion of achievement in education has been prevalent for some time (Crane, 1959). Expressed at its simplest, achievement implies the assessment of some level of attainment against a level expected. Under-achievement occurs when the level expected is greater than that attained; over-achievement is recognised when the level attained exceeds that anticipated. Central, then, to the notion of achievement is the identification of an acceptable and appropriate score against which attainment can be assessed. In reading the simplest *expected reading score* (ERS) is *chronological age* (CA), or stated more accurately *that score in reading achieved on average by pupils of that chronological age.* When this comparison is made the *attained reading score* (ARS) needs to be expressed as a reading age (RA); when RA exceeds CA over-achievement is recognised, when RA is less than CA under-achievement is recognised. Even though CA is a useful estimator of expected attainment, it fails to accommodate, totally, other variables which have been shown to correlate with reading, e.g., intellectual potential, social class, etc. In other words it is possible to calculate ERSs more accurately if the effect of other variables besides CA are included in the calculation.

Mental age, which is derived from CA and IQ, has been used (Critchley, 1964; Naidoo, 1970). Also formulas have been developed which not only account in part for potential for reading but also for the length of time a pupil has experienced a formal programme of reading instruction. Harker (1977) has summarised the rationale behind some of the formulae available, as well as highlighting their limitations. Mcleod (1979) and Shephard (1980) have also been critical of the formula approach for the identification of expected score. In general this criticism centres on the fact that the correlation between reading and other significant variables is not perfect and that regression to the mean is a statistical consequence of coefficients which are less than unity.

Use of a regression equation produces the most accurate estimate of ERS, in as much as it is the only approach which accounts, acceptably well, for the pheneomenon of regression to the mean (Shephard, 1980). The regression model directly operationalises the definition of achievement (under and over) as a discrepancy between attained and expected scores, where the expected score can be conceptualised as the *average score in the dependent variable (say reading) for subjects of that sample for all levels of the predictor variable(s)* (say CA, IQ, SES, etc.). As the regression line represents the average score in the dependent variable achieved for all levels of the predictor variable there will be approximately as many who underachieve as there will be those who overachieve.

A few small-scale studies in the UK have used this approach to identify readers for remedial treatment (Ravenette, 1961; Fransella and Garver, 1965). There has been one significant study (Rutter *et al.*, 1970) from which has emanated a number of papers including Berger *et al.* (1975), Rutter and Yule (1973, 1975), Yule (1973), and the most significant for this paper, Yule *et al.* (1974) who reported the prevalence of severe under-achievement (labelled specific reading retardation) to be greater than that predicted from the assumption of normally distributed residuals. Yule's "hump" was interpreted as evidence supporting the notion that "there is a group of children with severe and specific reading retardation which is *not* just the lower end of a normal continuum" (Yule *et al.*, 1974). This finding itself is significant as it differs from that implied in the Green Paper (DES, 1972) which recognises the existence of two qualitatively different groups of poor readers, namely

> (i) "children whose reading abilities are significantly below the standards which their general abilities in other spheres would lead one to expect"

and

> (ii) "children whose reading backwardness is only one aspect of their general retardation,"

but not in excess of the number anticipated by a normal distribution. Yule *et al.* (1974) were careful to point out that the existence of a "hump" does not necessarily imply any kind of pathological or biological identity. Nevertheless, its presumed existence has led to the interpretation that a group of readers do exist at the bottom tail of an achievement distribution in reading who are pathologically different from backward readers (Glossop *et al.*, 1979).

Rodgers (1983) has outlined a number of reasons supporting a need for caution in accepting the notion of Yule's "hump". His concern is with the characteristics of the data sets from which the existence of the "hump" was promulgated. They are: (i) the ceiling effect of the group reading test used resulted in a deviation from linearity of the regression functions, (ii) some degree of negative skewness in the histograms of the residuals, and (iii) the considerable variability in the incidences of those who present specific reading retardation across five samples. Rodgers implies that reasons (i) and (ii) are violations of assumptions underlying use of the regression procedure while (iii) is a manifestation of the violations.

Rodgers (1983) used data generated by the Child Health and Education Study to seek out Yule's "hump". He used a similar procedure to that followed by Yule *et al.* (1974), namely an intelligence score (four selected subtests from the British Ability Scales (Elliott *et al.*, 1979)) as the predictor variable on reading, which was assessed by a specially shortened version of the Edinburgh Reading Test (Godfrey Thomson Unit for Educational Research, 1977; Moray House College of Education, 1980). Age was controlled by sample selection. The sample size (N = 8836) was considerably greater than any of the samples of Yule *et al.* (1974). Rodgers concluded that "there is no discontinuity in the distribution of differences between obtained and predicted reading scores and indeed there is no suggestion of any bimodality which could be taken to indicate the presence of two underlying populations". In other words he failed to establish the existence of a "hump".

Silva *et al.* (1985) have partially replicated the Rutter and Yule (1975) study and in so doing commented on the visibility of the "hump". They used the same regression technique (eight subscales of the WISC (Weschler, 1974), the Burt Word Reading Test (Scottish Council for Educational Research, 1976)) and the same criterion for identification of "specific reading retardation", namely a difference greater than two standard errors between ARS and ERS. They found 1·2 per cent at the lower tail of the distribution of residuals rather than the 2·28 per cent anticipated from the curve of normal distribution — a "concavity" rather than a "hump".

Van der Wissel and Zegers (1985) also question the existence of the "hump". Their general argument is much like that of Rodgers (1983) in that they consder the "humps" recognised by Yule *et al.* (1974) to be statistical artefacts created by the "ceiling" effect on the reading test. They also argue that differences in the ratio of boys to girls with decreasing levels of underachievement which were recognised by Yule *et al.* (1974) are explained by a difference in mean scores and variances of the two distributions, i.e., the mean reading score for boys is lower than that for girls and the variance of the boys' distribution is greater.

To date then, Yule *et al.* (1974) support the existence of the "hump", Rodgers (1983), Silva *et al.* (1985) and van der Wissel and Zegers (1985) do not. This paper also seeks a "hump".

### *Methodology and Results*

Children (N = 5214; age range 9·00-10·75 years, $\bar{X}$ = 9·70 years) in all J3 classes (N = 285) of a local education authority were given the NFER BD test of reading (NFER, 1971) and the Raven's Standard Progressive Matrices (Raven, 1956). Table 1 gives a summary of the descriptive statistics for the distribution of reading and non-verbal intelligence scores; it also gives the correlation between the distributions as well as selected regression statistics.

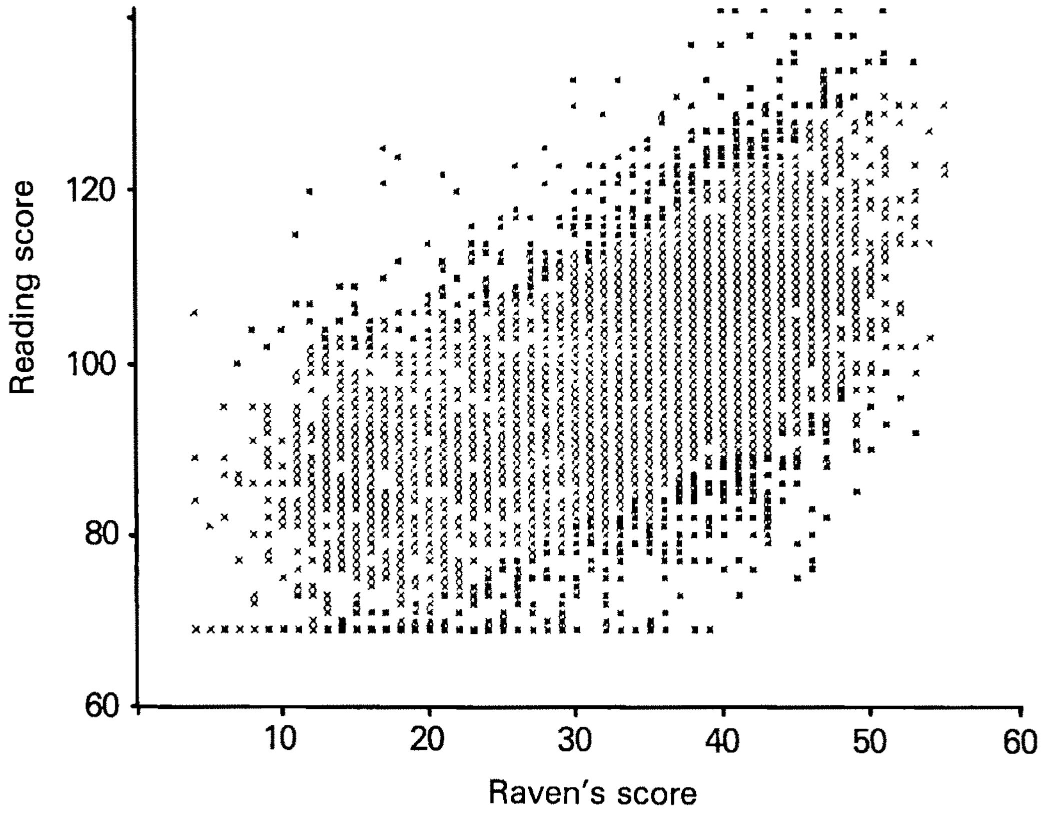
TABLE 1

Descriptive Statistics for the Distribution of Reading Scores (Reading Quotient Points) and Non-Verbal Intelligence (Raven's Raw Scores), Also Regression Statistics Between the Two Variables

| *1 Descriptive statistics* | | | |
|---|---|---|---|
| *Variables* | *N* | *Mean* | *SD* |
| Reading (NFER BD) | 5241 | 99·39 | 13·18 |
| Non-verb intelligence (Raven's) | 5198 | 33·36 | 9·83 |
| *2 Regression statistics* | | | |
| N = 5018 | Corr. = 0·61 | Slope = 0·77 | Const = 73·67 |
| Std. Error of Predict. = 10·74 | F value for regression = 2546·84 | | |

FIGURE 1

Scattergram of NFER BD Reading Scores (Reading Quotient Points) and Non-Verbal Intelligence Scores (Raven's Raw Scores)

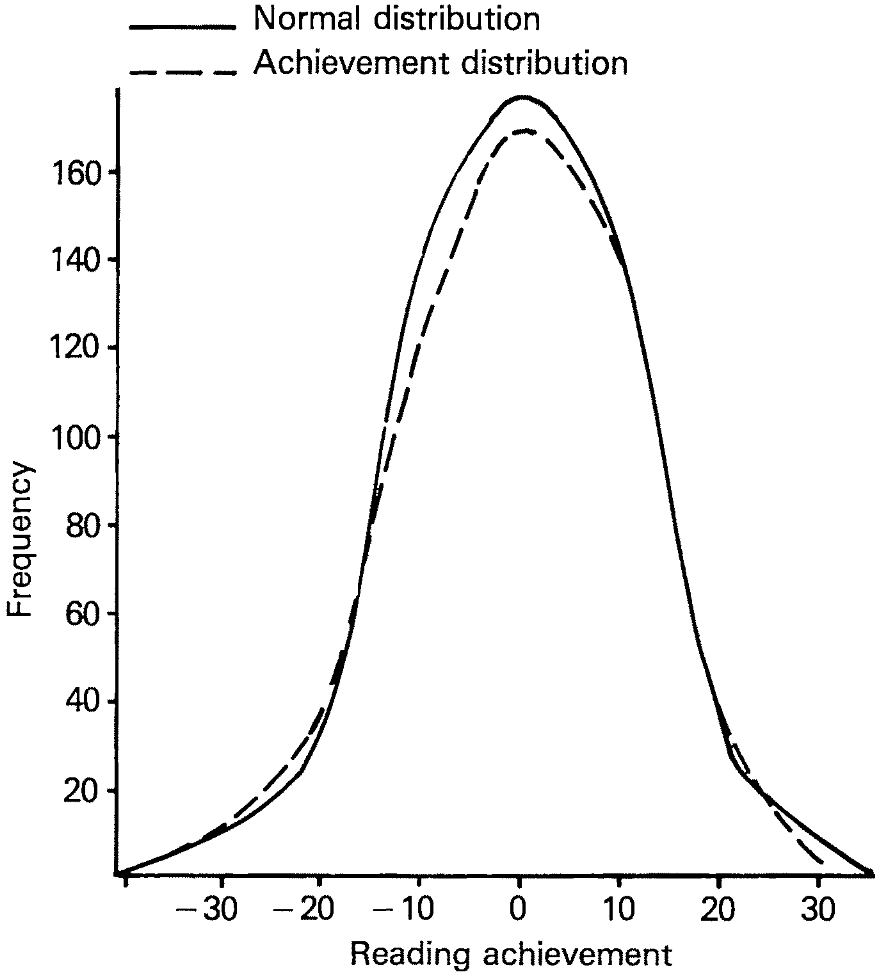


The distribution of reading scores was about that expected from the standardisation data (Mean = 100, SD = 15; NFER, 1971) and did not depart from normality when tested after the fashion of Pearson and Hartley (1966) for ungrouped data. The Raven's distribution also was much as expected and in line with the latest available norms from the Scottish Research Council (1981). The correlation of 0·61 reflects very closely that calculated by others, notably Rutter *et al.* (1970), between measures of reading and non-verbal intelligence. The substantive relationship between CA and attainment in reading over the age range of the primary school would, most typically, warrant the inclusion of CA in any regression equation to calculate an ERS. However, the restricted age range of the sample in conjunction with the use of the RQ metric, which is itself adjusted for age, resulted in a correlation between CA and ARS of 0·06. Consequently the regression was of the simple type with the Raven's scores acting as the single predictor variable. A scattergram of the two distributions (Figure 1) shows the criterion for linearity to be satisfied, also the lack of an hour glass or fan shape in the scores suggests no serious departure from the assumption of homoscedasticity.

What of the residuals of this study? Figure 2 shows a computer drawing of the smoothed

FIGURE 2

DISTRIBUTION OF READING ACHIEVEMENT SCORES (EXPECTED SCORE-ATTAINED SCORE) EXPRESSED AS READING QUOTIENT POINTS AGAINST THE EXPECTED NORMAL DISTRIBUTION



residual distribution against a normal distribution created with the same mean score (0·04) and standard error (10·74) as the distribution of residuals. Reading achievement is expressed in standard scores. Visually, at least, below two standard errors there is an increased frequency of subjects in excess of those predicted by the normal curve.

Table 2 is the numeric equivalent of Figure 2. It gives the actual frequency, also percentages, of subjects between selected standard error ranges. Below two standard errors the normal curve expectancy is 2·28 per cent of the population of the curve, or, in this case· 2·28 per cent of 5018 subjects, 114 subjects; the actual count is 118 — more of a "pimple" than a "hump" and not significant when tested for difference.

TABLE 2

ACTUAL AND EXPECTED FREQUENCIES OVER DIFFERENT REGIONS OF THE DISTRIBUTION OF READING ACHIEVEMENT SCORES (ARS-ERS)

READING ACHIEVEMENT IS EXPRESSED IN STANDARD ERRORS OF PREDICTION FROM THE MEAN

| | | Below mean | | Above mean | | |
|---|---|---|---|---|---|---|
| | 2·00 | 2·00-1·00 | 1·00-1·00 | 1·00-2·00 | 2·00 | Total |
| *Actual frequency* | | | | | | |
| | (135) | | | | | |
| Observed | 118 | 721 | 3326 | 729 | 124 | 5018 |
| Expected | 112 | 682 | 3425 | 682 | 112 | 5013 |
| *Per cent frequency* | | | | | | |
| | (2·69) | | | | | |
| Observed | 2·31 | 14·36 | 66·28 | 14·52 | 2·47 | 99·94 |
| Expected | 2·28 | 13·59 | 68·26 | 13·59 | 2·28 | 100·00 |

Adjusted frequencies in parentheses.

However, given the parameters of the regression equation with a constant of 73·67 and slope of 0·77, a Raven's score of 23 or below could not produce a difference greater than two standard errors between ARS and ERS. The lowest possible ARS is 70 reading quotient points. The ERS for a Raven's score of 23 is Constant + Slope (Raven's Score) or 73·67 + 0·77 (23) to give 91·38 or just less than the lowest possible ARS plus two standard errors (70 + 2 (10·74) = 91·48). Consequently the figure of 118 subjects must be regarded as an underestimate of the true number who might fall in the lower tail as a certain percentage of subjects are unable to register a difference of more than two standard errors between ARS and ERS.

An adjustment is possible. The Raven's distribution approximates the normal with a mean of 33·41 and a standard deviation of 9·78. A score of 23 is 1·06 standard deviation units from the mean score, a point below which there are some 725 subjects (N = 5018). If the assumption is made that the percentage of this group of 725 subjects who show an ARS more than two standard errors below the ERS is the same for the 4293 subjects (5018-725) who can present a two standard error discrepancy at 2·4 per cent, then the frequency of 118 children should be adjusted upward by 2·4 per cent of 725 subjects or by 17 subjects. Table 2 shows the adjusted count to be 135 subjects which is significant at the 0·05 level when tested against the number expected from the normal curve (chi square = 33·84, critical value (0·05) = 3·64).

However, this adjustment is made against a mean residual score of 0·04 RQ points and is for scores which will fall more than minus two standard errors from the mean. It necessarily follows that changes to scores on one side only of a distribution will contribute to a change in the mean score for that distribution. Consequently, the adjustment followed here would change the location of the mean score some (unknown) distance toward the lower tail. As a consequence of this relocation the score on the distribution which is two standard errors from the mean will also change.

Given the above argument, the adjustment calculated here must be seen as a partial correction as it accommodates only an increased frequency which is unbalanced by the relocation of the mean and its associated standard error. Re-calculating the mean score after adjustment to some scores in a distribution is difficult and a number of assumptions need to be made. Not least of these assumptions is that the lower tail of the distribution behaves in accord with "normal expectancy" and by so doing enables a judgment to be made on the basis of mathematical probability. Seemingly this assumption would not be acceptable to Yule *et al.* (1974).

A definitive answer cannot be given to the question of an increased frequency of underachievers who score beyond two standard errors below the mean. This increase recognised in three of the five distributions of Yule *et al.* (1974) is not recognised in the distribution presented here, statistically at least, when tested at the 0·05 level. However, it is recognised at the same significance level after application of the adjustment routine adopted by Yule *et al.* (1974).

If the preferred interpretation is to regard the adjusted frequency as evidence to support the notion of Yule's "hump" then it should be clearly recognised that the magnitude of the increased frequency over that expected is 23 children in a distribution of size greater than 5000.

Correspondence and requests for reprints should be addressed to Dr. D. A. Dobbins, Department of Education, University College of Swansea, Hendrefoilan, Swansea, SA2 7NB, Wales

## REFERENCES

BERGER, M., YULE, W., and RUTTER, M. (1975). Attainment and adjustment in two geographical areas: II The prevalence of specific reading retardation. *Br. J. Psychiat.*, 126, 510-519.

CRANE, A. R. (1959). An historical and critical account of the accomplishment quotient idea. *Br. J. educ. Psychol.*, 29, 252-259.

CRITCHLEY, M. (1964). *Development Dyslexia.* London: Heinemann.

DEPARTMENT OF EDUCATION AND SCIENCE (1972). *Children with Specific Reading Difficulties.* London: HMSO.

ELLIOTT, C. D., MURRAY, D. J., and PEARSON, L. S. (1979). *British Ability Scales Manual 4: Tables of Norms and Abilities.* Windsor: NFER.

FRANSELLA, F., and GARVER, D. (1965). Multiple regression equations for predicting reading age from chronological age and WISC verbal IQ. *Br. J. educ. Psychol.*, 35, 87-89.

GLOSSOP, J. A., APPLEYARD, R., and ROBERTS, C. (1979). Achievement relative to a measure of general intelligence. *Br. J. educ. Psychol.*, 49, 249-257.

GODFREY THOMPSON UNIT FOR EDUCATIONAL RESEARCH (1977). *Manual of Instructions for the Edinburgh Reading Test Stage 4.* Sevenoaks: Hodder and Stoughton Educational.

HARKER, W. J. (1977). How valid are reading potential scores?: *Remed. Educ.*, 12, 27-31.

MCLEOD, J. (1979). Educational underachievement: towards a defensible psychometric definition. *J. learn. Disabil.*, 12, 322-330.

MORAY HOUSE COLLEGE OF EDUCATION (1980). *Manual of Instructions for the Edinburgh Reading Test Stage 3.* Sevenoaks: Hodder and Stoughton Educational.

NAIDOO, S. (1970). The assessment of dyslexic children. In FRANKLIN, A., and NAIDOO, S. (Eds.), *Assessment and Teaching of Dyslexic Children.* Invalid Children's Aid Association.

NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH (1971). *Reading Test BD: Manual of Instructions.* Windsor: NFER — Nelson.

PEARSON, E. S., and HARTLEY, H. (Eds.) (1966). *Biometrikas Tables for Statisticians, Vol. 1.* Cambridge: Cambridge University Press.

RAVEN, J. C. (1956). *Standard Progressive Matrices.* London: H. K. Lewis.

RAVENETTE, A. T. (1961). An empirical approach to the assessment of reading retardation. *Br. J. educ. Psychol.*, 31, 96-102.

REITAN, R. M. (1972). Verbal problems solving as related to cerebral damage. *Percept. mot. Skills*, 34, 515-524.

RODGERS, B. (1983). The identification and prevalence of specific reading retardation. *Br. J. educ. Psychol.*, 53, 369-373.

RUTTER, M., TIZARD, J., and WHITMORE, K. (1970). *Education, Health and Behaviour.* London: Longman.

RUTTER, M., and YULE, W. (1973). Specific reading retardation. In MANN, L., and SABATINO, D. (Eds.), *The First Review of Special Education.* Buttonwood Farms: J.S.E. Press.

RUTTER, M., and YULE, W. (1975). The concept of specific reading retardation. *J. Child Psychol. Psychiat.*, 16, 181-197.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1976). *The Burt Reading Test, 1974 Revision.* London: Hodder and Stoughton.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1981). *Standardisation of the Standard Progressive Matrices*. Edinburgh: SCRE.

SILVA, P. A., MCGEE, R., and WILLIAMS, S. (1985). Some characteristics of 9-year-old boys with general reading backwardness or specific reading retardation. *J. Child Psychol. Psychiat.*, 26, 407-421.

SHEPHARD, L. (1980). An evaluation of the regression discrepancy method for identifying children with learning disabilities. *J. spec. Educ.*, 14, 90-91.

VAN DER WISSEL, A., and ZEGERS, F. E. (1985) Reading retardation revisited. *Br. J. dev. Psychol.*, 3, 3-9.

WESCHLER, D. (1974). *The Weschler Intelligence Scales for Children-revised.* New York: The Psychological Corporation.

YULE, W. (1973). Differential prognosis of reading backwardness and specific reading retardation. *J. educ. Psychol.*, 43, 244-248.

YULE, W., RUTTER, M., BERGER, M., and THOMPSON, J. (1974). Over and under-achievement in reading: distribution in the general population. *Br. J. educ. Psychol.*, 44, 1-12.

YULE, W., LANDSDOWN, R. and URBANOWITZ, M. (1982). Predicting educational attainment for WISC (R). *Br. J. clin. Psychol.*, 21, 43-47.

(*Manuscript received* 29*th October*, 1987)

*Br. J. educ. Psychol.*, 58, 345-349, 1988

# PREFERRED CLASSROOM ENVIRONMENT AND APPROACH TO LEARNING

By JOHN HATTIE
*(Department of Education, University of Western Australia)*
AND DAVID WATKINS
*(Education Department, University of Canterbury, New Zealand)*

SUMMARY. This study involved the administration of preferred classroom environment and approaches to learning measures to 1266 Australian high school students. Canonical analysis indicated a relationship between those learning strategies which are likely to lead to high quality learning outcomes and student preference for classrooms to be enjoyable and oriented to independent study and competition.

## INTRODUCTION

Classroom environment research has become a popular area of study in recent years (Moos, 1979; Fisher and Fraser, 1981; Haertel *et al.* 1981; Fraser and Fisher, 1983). The majority of these investigations have focused on the influence of the actual environment on achievement.

Much less research has examined the student's preferred classroom environment. However, an investigation by Hattie *et al.* (1986) has found that a generally preferred high school environment does exist but that this preference is moderated by age and gender. Grades 11 and 12 students wished to be involved in a cohesive network but at the same time wanted more self-initiated activities. Yet Grade 9 students were more inclined towards friction and competitiveness. Females showed a preference for social harmony but males were more in favour of competition. Hattie *et al.* argue that an apparent "decline" in classroom environment, reflected in lower actual environment scale means as students progress through school, is rather a change of preferences. Current classroom environment measures typically assess a limited number of dimensions often mixing them into one overall score. This makes it difficult to identify the change in preferences that may be occurring. Therefore, Hattie *et al.* advocate using an "economical" battery of instruments that tap various dimensions such as preferences for competitive, cohesive, independent, highly organised and teacher-managed classrooms.

Evidence of the relationship between the preferred classroom and learning outcomes is not strong. However, Fraser and Fisher (1983) found support for a person-environment fit hypothesis as the relationship between achievement and actual individualisation was more positive in classes where students tended to express a preference for individualisation. The research reported here utilises data collected in the Hattie *et al.* (1986) research to go beyond earlier studies by attempting to relate a student's preferred classroom environment to the way he/she studies.

Considerable advances have been made in research in this latter area in recent years. One of the most impressive things about these investigations has been the convergence of findings from very different research methodologies. Researchers utilising the "experiential perspective", typically involving intensive interviews with a small number of students to gain insights into how and why they go about their learning, have consistently shown that students tend to approach their learning in one of two ways — referred to as deep or surface approaches (Marton and Säljö, 1984). The deep approach is characterised by a student's intention to understand the material to be learnt plus the adoption of learning processes (such as inter-relating ideas, thinking about the structure as a whole) likely to lead to such understanding. On the other hand, the surface level approach is characterised by the student's intention to reproduce the material to be learnt and the adoption of rote learning strategies to achieve this.

Investigators utilising quantitative research methods have also identified and developed inventories to measure these two approaches to learning (Ramsden and Entwistle, 1981; Biggs, 1985). However, they have also added a third dimension involving achievement motivation and highly organised learning processes. Investigations of students' learning processes at tertiary level have shown that approach to learning depends on both contextual and personological factors (Ramsden and Entwistle, 1981; Watkins and Hattie, 1985). Thus, while each student tends to have a relatively stable stylistic preference for a particular approach to learning, their perception of particular learning contexts can influence them to adopt what they believe to be appropriate learning strategies. For example, a student whose normal style of learning is to achieve understanding by first gaining an overview and then building up supportive evidence (the characteristics of the comprehension learning style: Pask, 1976) may well adopt superficial learning strategies if faced by a multiple-choice examination; if little time is available to study; if the lecturer is perceived as providing little encouragement for independent thinking and/or teaches the course in a manner which arouses bordeom not interest (Watkins, 1984; Ramsden, 1985).

There is only limited evidence to date that the above environment/learning approach relationship holds at secondary school level (Selmes, 1986). However, it seems reasonable to propose that a student's perception of the *actual* classroom environment will be related to the *particular* strategy that is adopted therein, whereas the perceptions of the *preferred* environment is likely to be related more strongly to the *usual* approach to learning — itself a function of the preferred learning style. There is a major advantage to relating preferred environment to approach to learning rather than achieved grade. This is because a deep level approach has been shown to be necessary to achieve higher level learning outcomes such as critical thinking and independence of thought, but the latter are not often rewarded by academic grades (Watkins, 1984; Ramsden, 1985).

## METHOD

### *Subjects*

The sample consisted of 1266 students in Grades 7, 9, and 11 (aged 12-13, 14-15, and 16-17 years respectively) from 18 high schools on the north coast of New South Wales. The schools were located in small cities, suburbs, and rural areas and approximately equal numbers of boys and girls participated.

### *Instruments*

To measure preferred learning environments shortened versions of the Classroom Environment Scale (CES: Moos and Trickett, 1974), My Class Inventory (MCI: Fisher and Fraser, 1981), and the Individualised Classroom Environment Questionnaire (ICEQ: Fraser and Fisher, 1983) were utilised. The scales which constitute these inventories are shown in Table 1 together with the corresponding estimates of reliability (Alpha) as found in this research. The MCI contains 25 Yes-No format items, the CES 24 True-False format items, and the ICEQ 25 Likert format items (with 5-point scales from Almost Never to Very Often).

The students' approach to learning was assessed utilising the "How I study" (HIS) questionnaire — a shortened and modified version of the Study Process Questionnaire (SPQ: Biggs, 1985). The HIS, like the SPQ, has three scales, Surface Approach, Deep Approach, and Achieving Approach each of which has a motive and a strategy subscale. The subscales are described in Table 2 along with the corresponding estimates of reliability (Alpha) found in this research.

Each of the subscales consists of four items, selected from item analyses of the authors' previous use of the SPQ, but also with rewording of some of the original SPQ items. Each item is answered on a six-point scale from "1 = strongly disagree" to "6 = strongly agree".

### *Analysis*

The relationships between the preferred classroom environment scales and the study process scales were analysed by canonical correlation analysis. This analysis produces two important types of information — canonical variates and the canonical correlations between them. The canonical variates come in two sets. In our case, one for the preferred environment scales and the other for the HIS scales. The coefficients of these canonical variables represent the salience of these scales for the relevant canonical variate. The first canonical variate from

TABLE 1

SCALE DESCRIPTIONS FOR CLASSROOM ENVIRONMENT SCALE (CES), MY CLASS INVENTORY (MCI), AND INDIVIDUALISED CLASSROOM ENVIRONMENT QUESTIONNAIRE (ICEQ) AND OBTAINED ESTIMATES OF RELIABILITY (FOR SHORT FORMS)

| Scale | Description | Alpha Reliability |
|---|---|---|
| | *Classroom Environment Scale* | |
| Involvement | Students' attentiveness | 0·91 |
| Affiliation | Student' social awareness of each other | 0·93 |
| Teacher Support | Teacher's help and personal attention | 0·92 |
| Task Orientation | Amount of stress placed on curricular activity | 0·92 |
| Order and Organisation | Orderly behaviour and organisation of class | 0·90 |
| Rule Clarity | Students' knowledge of rules and consequences of breaking them | 0·92 |
| | *My Class Inventory* | |
| Satisfaction | Enjoyment of class | 0·92 |
| Friction | Aggressive behaviour of students | 0·89 |
| Competition | Importance of students' achievement relative to that of their peers | 0·92 |
| Difficulty | Difficulty of class work | 0·84 |
| Cohesiveness | Friendly relationship between students | 0·93 |
| | *Individualised Classroom Environment Questionnaire* | |
| Personalisation | Teacher's helpfulness and personal attention | 0·95 |
| Participation | Students' input in class work or discussions | 0·95 |
| Independence | Students' control over their work habits and general behaviour | 0·92 |
| Investigation | Students' own research | 0·93 |
| Differentiation | Students' freedom to work at their own pace and in their own style | 0·88 |

TABLE 2

DESCRIPTIONS AND OBTAINED ESTIMATES OF RELIABILITY FOR SCALES OF "HOW I STUDY" QUESTIONNAIRE

| Scale | Description | Alpha Reliability |
|---|---|---|
| | *Surface Approach* | |
| Surface Motive | Motivation is utilitarian; main aim is to gain qualifications at minimum allowable standard | 0·51 |
| Surface Strategy | Strategy is to reproduce bare essentials using rote learning | 0·61 |
| | *Deep Approach* | |
| Deep Motive | Motivation is interest in subject and its related areas | 0·71 |
| Deep Strategy | Strategy is to understand what is to be learnt through inter-relating ideas and reading widely | 0·66 |
| | *Achieving Approach* | |
| Achievement Motive | Motivation is to obtain highest possible grades | 0·71 |
| Achievement Strategy | Strategy is highly organised and designed to achieve high marks by being a "model" student, e.g., being punctual, doing readings, etc. | 0·68 |

each subset is chosen to maximally correlate with each other. Similarly for the second and all succeeding pairs of canonical variates. Individual student responses were used as the unit of analysis as seems appropriate when the preferred learning environment is concerned. After all, the latter is not necessarily the same for each class member (Hattie *et al.*, 1986).

## RESULTS

Three significant ($P < 0{\cdot}01$) canonical correlations were found between these sets of variables. However, as the third was below 0·30 in magnitude and provided no clear interpretation only the first two canonical variates will be discussed here (the coefficients for these are shown in Table 3).

TABLE 3

Coefficients of Canonical Variables Linking Preferred Classroom Environment Scales (Set I) with "How I Study" Scales (Set II). Only Loadings Above 0·29 are listed

| | Coefficients | |
|---|---|---|
| | I | II |
| *Set I* | | |
| Involvement | — | 0·36 |
| Affiliation | — | — |
| Teacher Support | — | — |
| Task Orientation | 0·30 | — |
| Order and Organisation | 0·31 | — |
| Rule Clarity | — | −0·38 |
| Satisfaction | 0·35 | 0·58 |
| Friction | — | — |
| Competition | — | 0·30 |
| Difficulty | — | — |
| Cohesiveness | — | −0·36 |
| Personalisation | — | — |
| Participation | — | — |
| Independence | −0·46 | — |
| Investigation | — | 0·54 |
| Differentiation | — | 0·44 |
| *Set II* | | |
| Surface Motive | — | −0·35 |
| Surface Strategy | — | −0·84 |
| Deep Motive | 0·66 | 0·37 |
| Deep Strategy | — | 0·52 |
| Achieving Motive | — | 0·46 |
| Achieving Strategy | 0·36 | −0·41 |

From Table 3, it can be seen that the first canonical variable ($R_c = 0{\cdot}46$) linked preferred classrooms perceived as being enjoyable, orderly, and stressing curricular activities but discouraging students working independently to an approach to learning characterised by interest in the subject and highly organised learning strategies. The second canonical variable ($R_c = 0{\cdot}33$) linked preferred classrooms perceived as being enjoyable and encouraging students to do their own research, to work at their own pace, and being involved in activities but within a competitive framework where neither rules nor friendship between students are over-emphasised, with an approach to learning characterised by deep rather than superficial or highly organised learning strategies and motivation to learn out of interest and/or for achievement rather than utilitarian purposes.

## CONCLUSIONS

This study has extended previous knowledge about the preferred classroom by linking it to the students' approach to learning. It indicates that deep level learning strategies, which are conducive to high quality learning outcomes, are reported by students who prefer classrooms to be enjoyable and oriented to students being encouraged to do their own research rather than relying on the teacher or a text-book and being allowed to work at their own pace and to choose their own way of learning. The latter includes the teaching being sufficiently individualised to allow different students to use different books and resource material and the teacher utilising a variety of teaching aids (e.g., blackboard, overhead projector, video, etc.) But such a classroom should retain a slightly competitive rather than socially oriented atmosphere. A satisfying but more teacher-oriented classroom is preferred by students who are interested in the subject but more likely to use highly organised learning strategies.

While it is always dangerous to infer causality from such data, the findings are in line with meta-cognitive theories that deep level learning strategies are more likely to be adopted by students who feel responsible for their own learning and who receive encouragement from their teachers to do this (Biggs, 1985; McCombs, 1986).

Correspondence and requests for reprints should be addressed to Dr. D. A. Watkins, Education Department, University of Canterbury, Christchurch 1, New Zealand.

## REFERENCES

BIGGS, J. B. (1985). The role of meta-learning in study processes. *Br. J. educ. Psychol.*, 55, 185-212.

FISHER, D. L., and FRASER, B. J. (1981). Validity and use of the My Class Inventory. *Science Educ.*, 65, 145-156.

FRASER, B. J., and FISHER, D. L. (1983). Student achievement as a function of person environment fit: a regression surface analysis. *Br J. educ. Psychol.*, 53, 89-99.

HAERTEL, G. D., WALBERG, H. J., and HAERTEL, E. H. (1981). Socio-psychological environments and learning: a quantitative synthesis. *Br. educ. Res. J.*, 7, 27-36.

HATTIE, J. H., BYRNE, D., and FRASER, B. J. (1986). Research into students' perceptions of preferred and actual learning environment. Paper presented at Annual Meeting of American Educational Research Association, San Francisco.

MARTON, F., and SALJÖ, R. (1984). Approaches to learning. In MARTON, F., HOUNSELL, D., and ENTWISTLE, N. (Eds.), *The Experience of Learning*. Edinburgh: Scottish Academic Press.

MCCOMBS, B. L. (1986). The role of the self-system in self-regulated learning. Paper presented at Annual Meeting of American Educational Research Association, San Francisco.

MOOS, R. H. (1979). *Evaluation Educational Environments: Procedures, Measures, Findings and Policy Implications*. San Francisco: Jossey-Bass.

MOOS, R. H., and TRICKETT, E. J. (1974). *Classroom Environment Scale Manual*. Palo Alto: Consulting Psychologist Press.

PASK, G. (1976). Styles and strategies of learning. *Br. J. educ. Psychol.*, 46, 128-148.

RAMSDEN, P. (1985). Student learning: retrospect and prospect. *Higher educ. Res. Dev.*, 4, 51-69.

RAMSDEN, P., and ENTWISTLE, N. J. (1981). Effects of academic departments on students' approaches to studying. *Br. J. educ. Psychol.*, 51, 368-383.

SELMES, I. P. (1986). Approaches to normal learning tasks adopted by senior secondary school pupils. *Br. educ. Res. J.*, 12, 15-27.

WATKINS, D. (1984). Student perceptions of factors influencing tertiary learning. *Higher educ. Res. Dev.*, 3, 33-50.

WATKINS, D., and HATTIE, J. (1986). A longitudinal study of the approaches to learning of Australian tertiary students. *Hum. Learn.*, 4, 127-141.

(*Manuscript received* 10*th November* 1987)

*Br. J. educ. Psychol.*, 58, 350-355, 1988

# SEX DIFFERENCES IN ATTITUDES TOWARD SCHOOL: STUDENT SELF-REPORTS AND TEACHER PERCEPTIONS

By EFRAIM DAROM AND YISRAEL RICH
*(School of Education, Bar Ilan University, Israel)*

SUMMARY. This research compares teachers' perceptions of boys' and girls' attitudes to school with the students' own reports. Students (N = 2670) in Israeli elementary, middle and high school classes (N = 77) reported on their general satisfaction, commitment to schoolwork and relations with teachers and their teachers evaluated each student's attitudes for these domains. Results indicated that in most classes girls have more positive attitudes than boys and that teachers overestimate the magnitude of the difference. Also, as students progress in school the power of student sex weakens as a predictor of teacher perception of student attitudes. Finally, there is evidence that characteristics of the school environment differentially affect teacher perception of boys' and girls' attitudes to school.

## INTRODUCTION

Numerous studies in several countries have investigated differences between boys' and girls' satisfaction with school and other school-related attitudes in a wide variety of educational settings. Excluding investigations of attitudes toward specific school subjects, most studies indicate that girls are more satisfied with school life and that their school-related attitudes are more positive than boys' (Berk *et al.*, 1970; Barker Lunn, 1972; Isherwood *et al.*, 1979; Gnagey, 1981; Rich and Darom, 1981). Further, a large body of research indicates that in general, elementary school teachers evaluate the school related behaviours of girls more positively than those of boys (Hartley, 1978; Medway and Egelson, 1980; Clifton *et al.*, 1986).

It could be argued that teachers may perceive greater sex differences in student attitudes to school than those reported by the children themselves. A number of researchers (see Clifton *et al.*, 1986) have demonstrated that many teachers maintain sex stereotyped expectations for student academic and personal behaviour. Although these expectations are probably based on real differences, they may become enhanced because of a tendency to overestimate the frequency of behaviours confirming stereotypes. This is due to the fact that confirmatory behaviours are more easily noticed and retrieved from memory than behaviour contradicting the stereotype (Hamilton and Gifford, 1976). Since there is also a tendency to view others' behaviour as dispositionally caused while ignoring the possibility of situational causation (e.g., Ross, 1977), it is likely that teachers will interpret boys' negative and girls' positive school behaviour as related to their attitude to school. Thus the greater openness of teachers to sex-stereotyped behaviours should lead to perception of magnified sex differences in school-related attitudes.

Accordingly, this study investigated three hypotheses:

1. In most classes girls report more positive attitudes to school as compared to boys.

2. Teachers perceive girls' attitudes to school as more positive than those of boys in most classes.

3. Teachers perceive greater sex differences for school-related attitudes as compared to those reported by students.

## METHOD

### *Sample*

Participants were 2670 students in 77 classes and their teachers, sampled from a broad array of Israeli elementary, middle and high schools that took part in a nation-wide survey of perceptions of the quality of school life (Darom and Rich, 1981).

Small (15 students or less) and gender imbalanced (more than 70 per cent of students the same sex) classes were excluded from the original sample to reduce the impact of outliers and confounding variables. The remaining 53 classes included 11 fourth grades, 10 fifth grades, 10 sixth grades, 6 seventh grades, 6 eighth grades, 4 ninth grades, 3 tenth grades, 1 eleventh grade and 2 twelfth grade classes.

*Measures*

The categories of student attitudes investigated were general satisfaction with school (SAT), commitment to schoolwork (COM), and relations with teachers (TCH). These three were measured by two methods using students' self ratings and teachers' ratings of students.

1. The Israeli Quality of School Life (IQSL) questionnaire (Darom and Rich, 1981), a student self-report 27-item instrument which is an adaptation of the Quality of School Life scale developed by Epstein and McPartland (1976) was used. Students' scores were defined as the total number of agreements to positive items and disagreements to negative statements regarding each of the three attitude categories. Examples of items from each scale are as follows: "I like school very much" (SAT); "I enjoy the work I do in class" (SWK); "Most of my teachers really listen to what I say" (TCH). Alpha reliabilities of these scales were 0·71, 0·76 and 0·68, respectively.

2. Teachers evaluated each of the students in their classes with respect to SAT, COM, and TCH on a 4-point scale as part of the effort to validate the IQSL. Teachers were not aware of students' responses.

*Procedure*

The IQSL was administered during class sessions in the absence of their teachers. At the same time, teachers were asked to evaluate each student in their classes. Both groups were informed that their anonymity would be kept, and that their responses would be used for research purposes only.

*Analysis*

The class was the primary unit of analysis because of the focus on teachers' evaluations. For some analyses the number of classes was quite small, so the individual student became the preferred unit of analysis.

At the class level of analysis the correlation between each of the six (three categories of attitudes by student self-reports and teacher evaluations) measures and the dummy coded sex variable was used as a measure of sex differences. Girls were coded "2" and boys were coded "1", hence a positive correlation indicates that girls have more positive attitudes towards school than do boys.

Non-parametric statistical tests were directed at determining the proportion of classes that exceeded certain standards. For example, the sign test was used to determine whether teacher perceptions of sex differences was greater than that of student self-reports in more than 50 per cent of the classes.

For the individual level of analysis multiple regression was employed in which the criterion (Y) was teacher perception of SAT, COM and TCH, while student sex and self-reported attitudes were the predictors. Teachers' mean evaluation and class' mean attitude were also included in each regression model to control for different baselines.

The unique variance shared by student sex and attitudes with each of the three teacher perception criteria was computed by obtaining the difference between the $R^2$s of two regression models. One model included all predictors while the other included all but sex or attitudes respectively. The analyses were carried out for each of the three domains of school life and for three age groups: fourth to sixth grades; seventh to ninth grades; and tenth to twelfth grades. The advisability of breaking down children's attitudes to school by age groups is indicated in earlier research (e.g., Rich and Darom, 1981) which demonstrated a general decline in students' scores over the years in school.

## RESULTS

The first row of Table 1 reports the class-level median correlation coefficient between student sex and students' self-reported attitudes towards three domains of school life. Note that these correlations are low and not significant. However, the percentage of classes with negative correlations as indicated in the second row of the table is significantly smaller than 50. The magnitude of the difference between the sexes ranged from 0·12 to 0·28 standard deviations for the three attitude scales. These data suggest that in most classes girls report more positive attitudes compared to boys but the difference between the sexes is not large. These results at the class level of analysis support previously reported results (Rich and Darom, 1981) at the student level of analysis.

TABLE 1

SEX DIFFERENCES IN ATTITUDES TOWARD SCHOOL FOR STUDENT SELF-REPORTS AND TEACHER PERCEPTIONS (N = 53 CLASSES)

| | | General Satisfaction (SAT) | Commitment to Schoolwork (COM) | Relations with Teachers (TCH) |
|---|---|---|---|---|
| Student Self-reports | Median correlation between student sex and student attitudes r (s) | 0·041 | 0·051 | 0·051 |
| | Proportion of classes with negative r (s) | 0·370* | 0·380* | 0·315** |
| Teacher Perceptions | Median correlation between student sex and teacher perception r (t) | 0·145 | 0·220 | 0·213 |
| | Proportion of classes with negative r (t) | 0·220** | 0·150** | 0·185** |
| | Proportion of r (t) exceeding r (s) | 0·570 | 0·780** | 0·680** |

* P < 0·05 ** P < 0·01

Analysis of the relationship between student sex and teacher perceptions of students' attitudes to school yields more substantial correlations as can be seen in row 3 of Table 1. These indicate that in a typical classroom the teacher evaluates girls' school-related attitudes as more positive than those of boys. Also, the proportion of classes with negative correlations between student sex and teachers' evaluations of attitude was smaller (see line 4 of Table 1) than that relating student sex and students' self-reported attitudes. The magnitude of differences ranged from 0·39 to 0·59 standard deviation.

These results indicate that that teacher perceptions of students' attitudes to school yield greater differences between boys and girls than do the children's self reports. A direct within-class comparison of the percentage of negative correlations for student sex and student self-reported attitudes versus that for student sex and teacher evaluations of students' attitudes generally confirms this result. As can be seen in the final row of Table 1, the proportion of classes in which teachers' evaluations as compared to student self-reports indicate more positive attitudes for girls than for boys is significantly larger than 50 for the COM and TCH scales.

To explore further the relationship between student sex and teacher perceptions of student attitudes to school, students were broken down into three age groups — fourth to sixth grades, seventh to ninth grades and tenth through twelfth grades. The focus of interest was on determining whether teacher evaluations of students' attitudes are related to student sex throughout the school years or primarily in the early years when attitudes are most positive (see Rich and Darom, 1981). Multiple regression was employed as above, except that these analyses were conducted at the individual student level because of the small number of high school classes.

TABLE 2

INCREMENTAL VALIDITY ($R^2$ CHANGE) OF STUDENT SEX AND STUDENT ATTITUDES ON TEACHER PERCEPTION OF ATTITUDES

| Predictor | General Satisfaction (SAT) $R^2$ | Commitment to Schoolwork (COM) $R^2$ | Relations with Teachers (TCH) $R^2$ |
|---|---|---|---|
| *4th-6th (N = 941)* | | | |
| Sex | 2·2** | 3·5** | 3·3** |
| Attitudes | 0·4 | 2·5** | 0·7* |
| *7th-9th (N = 668)* | | | |
| Sex | 0·6 | 1·8** | 1·1** |
| Attitudes | 0·3 | 2·6** | 1·5** |
| *10th-12th (N = 190)* | | | |
| Sex | 0·1 | 0·7 | 0·0 |
| Attitudes | 1·9 | 5·0** | 0·9 |
| Results for 7th to 9th grades broken by type of school | | | |
| *Elementary (N = 213)* | | | |
| Sex | 1·0 | 2·3* | 3·2* |
| Attitudes | 0·5 | 0·0 | 0·8 |
| *Jr. High (N = 941)* | | | |
| Sex | 0·7 | 1·6* | 0·6 |
| Attitudes | 1·0 | 5·7** | 2·1* |

* $P < 0{\cdot}05$ ** $P < 0{\cdot}01$

Results of the multiple regression analyses, presented in Table 2, indicate that the sex of fourth to six graders adds significantly ($P < 0{\cdot}01$) to the prediction of teacher perceptions of all three domains of students' attitudes. Employing the same criterion for significance, student self-reported attitudes related to teacher perceptions of student COM only.

The pattern of results for the seventh to ninth grade pupils is quite different. Neither student sex nor student attitudes related to teacher perceptions of students' SAT. However, both predictors related significantly to teacher perceptions of students' COM and TCH.

At the high school level, student sex had no significant predictive value for any of the teacher perception criteria, while student attitudes contributed significantly to the prediction of teacher perceptions of student COM only.

These results indicate that student self-reported attitudes are related to teacher perceptions of student commitment to schoolwork at all grade levels. Student sex is related to teachers' evaluations of all school attitude domains investigated in the elementary school, to two of the three domains in the middle school range and to none of the teacher evaluation criteria in high school. This trend of diminishing relationship between student sex and teacher evaluations as children progress in age led to the question whether we are confronting a developmental phenomenon of maturation or the effects of different types of school frameworks affecting children and teachers in a differential manner (see Rich and Darom, 1981).

To provide a tentative test of the issue, students in the seventh, eighth and ninth grades were divided into two groups according to the type of school they attended. One group included children in the highest classes of elementary school while the other group was composed of students in junior high school, the vast majority of which were physically and administratively linked to senior high schools and shared a common teaching staff.

Results of this analysis, presented in the bottom section of Table 2, indicate that student sex is a better predictor of teacher evaluations of student attitudes in elementary school as compared to junior high school, while self-reported student attitudes in junior high school are better predictors of teacher evaluations than they are in elementary school. It is noteworthy that the greatest $R^2$ change appears for the attitude predictor on teacher evaluation of COM for students in junior high school and senior high school. Hence, these results provide tentative evidence that the trend of diminishing relationship between student sex and teacher's evaluation is not merely a maturation phenomenon, but is affected by the characteristics of the school attended.

## DISCUSSION

Results from this study confirmed the three hypotheses raised earlier. First, girls' self-reported attitudes to school were more positive than those of boys in most classes investigated. Second, teachers perceived that girls' attitudes were more positive than those of boys in over 75 per cent of the classes. Finally, in most cases, teachers perceived larger sex differences for commitment to schoolwork, and for relations with teachers than as expressed by the children themselves.

Although this study did not directly investigate why teachers perceive exaggerated sex differences in attitudes to school, the literature on the development of stereotypes (e.g., Hamilton and Gifford, 1976) and on the self-fulfilling prophecy (e.g., Darley and Fazio, 1980) provide a suitable theoretical rationale.

Noteworthy here, however, is the finding that teacher perceptions were most similar to those of students' self-reports for SAT while greater discrepancies were found for COM and TCH. Why should teachers' perceptions of sex differences for greater satisfaction be relatively accurate while they exaggerate sex differences in the other two domains?

It could be argued that teachers perceive that boys generally enjoy school as much as do girls. However, boys' satisfaction with school is not attributed to schoolwork but rather to camaraderie, sports and fooling around while girls' enjoyment of school is attributed to these as well as to schoolwork. Recall that boys were described as more impatient, lazier, noisier and less able to concentrate than were girls (Hartley, 1978). These behaviours are incongruent with teacher perceptions of commitment to schoolwork, but they do not rule out the perception of general satisfaction.

To explain teachers' perception of exaggerated sex differences for TCH we turn to the concept of "problem ownership" as conceptualised by Brophy and Rohrkemper (1981). They categorised student behaviour problems as teacher-owned, student-owned or teacher-student owned. Teacher-owned student behaviour problems, such as hostile aggression, under-achievement and defiance, prevent the teacher from meeting his or her needs or cause teacher frustration or anger. These kinds of behaviour are more typical of boys than of girls. Brophy and Rohrkemper demonstrated that teachers believe that students with teacher-owned problems act intentionally and should be able to control their behaviour. Teachers tend to attribute the causes of the behaviour to dispositional rather than situational sources. Since these obnoxious behaviours negatively affect teacher self-fulfilment, teachers are likely to attribute them to students' negative attitudes towards their relations with teachers. Hence, the preponderance of teacher-owned problems among boys as compared to girls should lead teachers to perceive greater sex differences for relations with teachers than are reported by students.

Another finding here that requires explanation is the trend of diminishing effect for student sex as a predictor of teacher perceptions of student attitudes. It should be noted that compulsory education in Israel ends at tenth grade. Thus, boys remaining after tenth grade may be more conforming to school norms and may display less negative behaviour than younger pupils. If this were the case, fewer differences between boys and girls would be evident in high school than in elementary school. Accordingly we would expect that elementary school student sex would have greater impact on teacher perceptions of student school-related attitudes than would student sex in high school. Although no direct evidence from this study is available to confirm this explanation, the increased predictive value of student attitudes on teacher perceptions of student commitment to schoolwork in high school is congruent with the rationale offered here.

Finally, attention is called to the finding regarding differences between similarly aged pupils in elementary versus junior high school. Earlier, it was argued that this phenomenon is probably related to characteristics of the two types of schools. The Israeli junior high school is usually linked physically and administratively to an academic or a comprehensive high school and the two generally have a shared teaching staff and common school norms (Rich *et al.*, 1981). Thus, it is not surprising that student sex loses its predictive value for teacher perceptions of student attitudes in junior high school, as is the case for tenth-twelfth grades, and student attitudes become more powerful predictors. On the other hand, the relationship for seventh and eighth graders in the same elementary school does not differ greatly from that for fourth to sixth graders. Although they have matured somewhat and the boys may be less rowdy than they were, from the vantage point of the seventh grade elementary school teacher the perception of sex differences in behaviour has become firmly entrenched. Expectations for the future may take into account the improved behaviour of boys, but drastic changes in teacher perceptions should not be anticipated. Accordingly, student sex retains its predictive power in the elementary school, but not in the junior high school.

Correspondence and requests for reprints should be addressed to Dr. Yisrael Rich, School of Education, Bar Ilan University, 52100 Ramat-Gan, Israel.

## REFERENCES

BARKER LUNN, J. (1972). The influence of sex, achievement level and social class on junior high school children's attitudes. *Br. J. educ. Psychol.*, 42, 70-74.

BERK, L., ROSE, M., and STEWART, D. (1970). Attitudes of English and American children toward their school experience. *J. educ. Psychol.*, 61, 33-40.

BROPHY, J., and ROHRKEMPER, W. (1981). The influence of problem ownership on teachers' perception of and strategies for coping with problem students. *J. educ. Psychol.*, 73, 295-311.

CLIFTON, R., PERRY, R., PARSONSON, K., and HRYNUIK, D. (1986). Effects of ethnicity and sex on teachers' expectations of junior high school students. *Sociol. Educ.*, 59, 58-67.

DARLEY, J , and FAZIO, R. (1980). Expectancy confirmation processes arising in the social interaction sequence. *Am. Psychol*, 35, 867-881.

DAROM, E., and RICH, Y. (1981). Development and validation of the Israeli Quality of School Life scale. In EPSTEIN, J. (Ed.), *The Quality of School Life.* Lexington: Lexington Books.

EPSTEIN, J., and MCPARTLAND, J. (1976). The concept and measurement of quality of school life. *Am. ed. Res J.*, 50, 13-30.

GNAGEY, W. (1981). Attitudes, values and motives of classroom facilitators and inhibitors. In EPSTEIN, J. (Ed)., *The Quality of School Life.* Lexington: Lexington Books.

HAMILTON, D., and GIFFORD, R. (1976). Illusory correlation in interpersonal perception: a cognitive basis of stereotype judgments *J. exp. soc. Psychol*, 12, 392-407.

HARTLEY, D. (1978). Teachers' definitions of boys and girls: some consequences. *Res. in Educ*, 20, 23-35.

HERZ-LAZAROWITZ, R., and SHARAN, S. (1979). Self-esteem, locus of control and children's perception of classroom social climate: a developmental perspective. *Contemp. educ. Psychol.*, 4, 154-161.

ISHERWOOD, G., AHOLA, J., HAMMAH, C., and SULLIVAN, P. (1979). *The Quality of School Life: Research Monograph.* Reports in Education No. 7, McGill University, Faculty of Education, Montreal.

MEDWAY, F., and EGELSON, R. (1980). Teacher ratings of internal and external students in open and conventional class environment. *Psychol. in Schools*, 17, 390-395.

RICH, Y , AMIR, Y., and BEN-ARI, R. (1981). Social and emotional problems associated with integration in Israeli junior high schools. *Int. intercult. Relat*, 5, 259-275.

RICH, Y., and DAROM, E. (1981). Pupils' perceptions of the quality of life in advantaged and disadvantaged Israeli schools. In EPSTEIN, J. (Ed.), *The Quality of School Life.* Lexington: Lexington Books.

ROSS, L. (1977) The intuitive psychologist and his shortcomings: distortions in the attribution process. In BERKOWITZ, L (Ed.), *Advances in Experimental Social Psychology. Vol. 10.* New York: Academic Press.

(*Manuscript received 18th August,* 1987)

*Br. J. educ. Psychol.*, 356-360, 1988

# ESSAY REVIEW

## DO THE GROUNDS FOR CLAIMING THAT SCHOOL MATTERS, MATTER?

BY DAVID MCNAMARA
*(School of Education, University of Durham)*

MORTIMORE, P., *et al.* (1988). *School Matters.* Wells, Somerset: Open Books, pp. 310, £7·95 pbk., ISBN 0-7291-0194-0.

### INTRODUCTION

Parents, teachers and sundry educationists are convinced by their experience that schools matter. Secondary heads often comment upon the substantial differences in the attainments of children from their feeder primary schools; parents will make the financial sacrifice to opt out of the state system because they consider that their children are not achieving their full potential; indeed, many of the provisions of the Reform Bill are based, in part, upon the premise that the public at large are concerned that schools matter.

In recent years there has been a substantial corpus of research which has sought to identify why some schools are more effective than others. Some of the findings which researchers claim to have discovered are now considered so well established that they can be packaged as tips for teachers as in *What Works* (Bennett, 1987) which has sold over half a million copies in the USA, even though it contents have not gone unchallenged (Glass, 1987).

The latest contribution to this enterprise is *School Matters: the Junior Years* (Mortimore *et al.*, 1988) which is destined to become a significant text which will be made much of by inspectors, advisers, teacher trainers and pundits. Upon publication the *Times Educational Supplement* devoted four pages to reviews and exposition and *Education* made it its document of the week.

*School Matters* is a substantial, well resourced, longitudinal survey which uses a variety of methods of data collection and sophisticated statistical analyses in order to investigate whether the junior school attended by a child makes a difference. To quote: "Will a child's progress in reading or writing be similar wherever she or he is taught? Are some schools more effective than others?" (p. 1). The authors begin with the confident assertion that "we researchers have managed to clear the decks, and to develop generally accepted ways of answering the questions with which we begin this chapter" (p. 2). This claim may come as news to scholars such as Shulman (1986), Colvin (1987), and Biddle and Dunkin (1987) who have effectively demonstrated the substantial conceptual and methodological problems attendant upon school effectiveness research. The certitude of *School Matters* is hardly matched by the tentativeness of Zirkel and Greenwood (1987) who note "that the far-flung prescriptive pronouncements currently in vogue, whether as recipes or as frameworks for school improvement, are not at all clearly justified by the research on effective schools" (p. 256).

In what follows I do not wish to discuss and evaluate the *School Matters* "findings"; indeed I would suggest that much of what the authors have to say is eminently reasonable. But the notion that the implications of this book for primary practice are based upon a secure empirical foundation must be challenged. We must approach *School Matters* with a measure of circumspection.

### THE SAMPLE

As the authors remind us, "There is strong evidence that schools vary considerably in the areas they serve and the background characteristics of their pupils" (p. 167), and that the Inner London Education Authority, the authority in which the research was carried out, is generous in its provision (p. 264). It is clearly the case that ILEA and its primary schools are atypical. For example, in 1985 67 per cent of ILEA schools had from 51 to 200 pupils while the national average was 52 per cent, and while 10 per cent of the nation's schools had fewer than 51 pupils there were only 0·4 per cent of ILEA schools classified in this bracket. At that time

ILEA allocated 58·3 per cent of its gross expenditure to teaching, while the national average was 67·8 per cent (CIPFA, 1984). Nearly half the schools in the sample were Victorian "three-deckers" and few had playing fields or open spaces. Forty per cent of the pupils were from black or other ethnic minority groups and 16 per cent of the children did not use English as their first language. At the start of the survey 31 per cent of pupils were receiving free school meals compared with a national average of 21 per cent.

These data describe primary schools in a distinctive environment, namely the deprived inner city area which serves the needs of the more disadvantaged and marginalised members of the community. It must be stressed that this is not the typical context in which primary education takes place. One must seriously question whether research undertaken within this context has general applicability. Indeed the authors recognise that the relevance of their findings might be greatest in the inner city areas (p. 262). Unfortunately this caveat seems to have been forgotten during the process whereby the general findings emerging from this research have been made available to the lay public. We must be cautious also when the authors seek to instruct us as to how we ought to train student teachers. A greater recognition on their part that primary education takes places in a variety of social and cultural environments within our society and that there is not and never can be one best way to teach all children may be in order.

## THE METHOD

Method, as conceived in this study, is to employ a longitudinal design, collect input, process, and output information, and analyse the resultant data set using multivariate statistical models. The methods employed have not been located within a coherent theoretical or conceptual framework which provides the context for the interpretation of the evidence. Indeed, as will be indicated below, because of the somewhat eclectic and atheoretical nature of the findings they can be invoked by the radical or conservative in order to defend an educational policy. The original design intended that 2,000 children would be tracked through 50 randomly selected primary schools during the four years of their junior careers. In the event, the initial sample consisted of 1,826 pupils and of these 1,411 remained in the same school during the first three years of their junior education. Presumably even fewer were tracked through for all four years. On average there would have been, say, about 30 children followed through in each school. Since the average number of classes in each school was eight we may presume that the 30 children split between two classes for each cohort, say 15 children per class. It is made quite clear that the unit of analysis in the design is the individual child; thus in this study we can only infer that we are not considering teacher or school effects upon coherent classroom groups of children but only upon randomly selected individuals within each class. As the authors indicate, this research design has to accommodate "severe limitations" (p. 7) and that "in order to overcome some of the methodological problems caused by these limitations a variety of statistical techniques were employed" (p. 7).

Two observations are in order. The first is that it may be appropriate to use these data in order to provide statistical descriptions of the circumstances in which children are educated in junior classrooms and to provide data about the changes wrought among statistical aggregates of pupils. But it is a logical fallacy to make inferences about classrooms and teachers. In order to study the interaction between teacher and pupils, and to make inferences about the complex process of education as it takes place within the real time constraints of classrooms, it is essential that the classroom is the unit of analysis. The second point is that one can endlessly debate the extent to which it is possible to deploy complex statistical models in order to overcome the difficulties of methodological design and the extent to which statistical methods can partial out the complexities of the human condition. But in the end one makes a commitment to a paradigm, and there is a sense in which it is no more than a value judgment to believe that the myriad effects of the richness, mundaneness, diversity and subtleties of the processes of primary education over a four-year period can be partialled out using statistical processes so as to reveal significant causes and effects.

## THE DATA

The authors locate their work within the research style epitomised by Bennett (1976) and Rutter *et al.* (1979) and a distinguishing characteristic of these texts is that they provide the reader with the data upon which their work is based. Moreover, precisely because the data are freely available, informed experts have subsequently been able to appraise critically the evidence, and indicate why they think it is deficient or needs to be treated with circumspection. To do this is nothing more than engage in normal scholarly practice. *School Matters*, by contrast, is virtually data free in the sense that the study is clearly based upon a substantial data set and various statistical analyses, yet the reader is not provided with the data so as to make an independent judgment of its worth and relate it to the written text. Since the authors claim significance for their work, the reader deserves better than to be presented with findings or evidence drawn from the data set in order, presumably, to substantiate what the authors take to be significant. Time and again one needs to be able to refer to the data while reading the text and it will not do for the authors continually to refer the reader to technical reports which are not freely available. Since Open Books could afford to publish the data for *Teaching Styles* and *15,000 Hours*, one wonders why they could not provide the same service in the case of *School Matters*.

It is not mere pedantry to make this point, because when data are presented there are many occasions when one has cause for concern. To take a representative example which relates to the effects which schools have on promoting pupils' reading scores. The text reads: "The size of the effects of particular schools upon children's reading progress in terms of raw score points was calculated and compared. Our figures show that the difference between the school with the most positive and that with the most negative was a striking 25 points (see Table 9.4)" (p. 187). When one refers to Table 9.4 the reader is simply provided with the data for the most effective and the least effective school, plus an overall average score. Anyone can make a dramatic point and indicate a significant effect by merely taking the two extreme points of a distribution. What is required are the data for the total sample and for the data to be presented in such a way that one can discover the extent to which all schools within the sample may or may not be effective on this dimension.

## AN ILLUSTRATION

*School Matters* is a long and densely packed book and one can only illustrate reservations by nominating examples. I will take the chapter on teacher expectations (p. 163ff.). There is a vast literature which addresses the conceptual, theoretical and methodological problems attendant upon researching teacher expectations; see, for example, Wineburg (1987) and Braun (1987). However, the essence of the notion of teacher expectations is that it is a mental construct and that teachers create "realities" which affect their behaviour towards pupils. After a quick review of some of the literature, the authors present us with their findings. Findings which clearly indicate that they focus upon teachers' actual behaviour rather than their mental dispositions. The chapter is peppered with claims such as: "Teachers also listen to autumn born pupils read more frequently than they listen to their younger peers even though the older children scored more highly on reading assessments and were rated as being of higher ability" (p. 164), no data provided; "Pupils whose fathers were not working read to their teachers more often than those who lived with only one parent" (p. 165), no data provided; and "Within the classroom it appeared that, in the main, teachers responded to the needs of lower ability pupils by giving them extra support with their work and hearing them read more often" (p. 171), no data provided. But the point is that the authors are clearly investigating teacher behaviour. Even though they do not provide the data, presumably we can, in all good faith, accept that the authors report accurately what the data show. However, later on they go on to say "low expectations may be transmitted in very subtle ways" (p. 171). This, of course, is precisely the problem and it is precisely the sort of problem that researchers investigating teacher expectations are supposed to tease out. It is illegitimate for the authors to go beyond their data and hazard an inference that low expectations may be transmitted in subtle ways, especially when the very stuff of teacher expectancy research is to try and elicit how such expectations may affect teacher pupil interaction.

More importantly, teacher expectations must be investigated within the context of the classroom; what is important is to discover how expectations may shape the teacher's behaviour towards children with differing attributes within her class. In this study one can only assume that the data relate to the sample children in each class and that the data are aggregated for the whole sample. Moreover, one does not know how the data relate to children's progression through the four years of junior education. Are children with unemployed fathers at a disadvantage when compared with children from one parent families during all years of junior schooling and in all classrooms, and if this is the case where are the data demonstrating the statistical significance of this finding, and what might be its educational significance?

## UNDERSTANDING EFFECTIVENESS

Perhaps the section of the book which has attracted must public attention is that part which claims to establish the characteristics of effective schools. One could spend much time discussing whether or not the so called characteristics of effective schools are correlates of effectiveness or causes of effectiveness. Some 28 variables are tested in order to elicit the characteristics of effective schools, ranging from the nature of the buildings, the size of the school, the head's style of leadership, the rewards and punishments used, teacher characteristics and parent involvement, to name but some of the somewhat long and eclectic list. The effects of these variables were tested, presumably by various statistical analyses, the details of which, unfortunately, are not provided. It is reported for instance (p. 251) that the deputy head can have a major role to play in promoting the effectiveness of junior schools. Where the deputy was frequently absent, or absent for prolonged periods, this was detrimental to pupils' progress. One yearns to know how this finding, which is not located in any theoretical or substantive context, has educational relevance. One can speculate why the role of the deputy head may be important and in order to engage in such speculation one has to draw upon one's common experience of primary schooling. The precise way in which the absence of the deputy head might affect student learning can only be discussed and elaborated and illuminated within the context of a study which conceptualises the school as a coherent unit. As it is, one wonders why it might be that the presence or absence of the deputy head has some sort of effect upon pupils' cognitive development and learning.

Some 25 years ago Rabinowitz and Travers (1953) reminded us that effectiveness cannot be discovered by empirical or statistical means; it is something which is imposed from without. Any attempt to investigate the effectiveness of schools involves a value judgment in terms of deciding which variables to nominate for investigation and how they are to be measured. More specifically Shulman (1986) observes that "It should be clear that no use of correlative criteria can be free normative choices".

In a word, the very ways in which schooling could be seen as effective in *School Matters* were built into the study from the outset, but the theoretical rationale for their inclusion is not convincingly defended.

## IMPLICATIONS OF THE RESEARCH

Since this research was funded and supported by ILEA it is understandable that the authors provide an interpretation and account of their findings which support the continuing provision of primary education with a local authority context and they indicate ways in which primary education could be enhanced within the framework of the 1944 Education Act. But while it may be a little mischievous, it has to be pointed out that the central findings emerging from this study will offer greater support and comfort to those who advocate the provisions contained within the Reform Bill. Indeed, it has to be noted that the research team could only discover what they claim to be effective schools via a massive investment in testing. Recently Sir Rhodes Boyson has used this research when addressing the Freedom Association in support of his proposals to privatise the examinations system (*Education*, 1988). This is the fate of all eclectic research; its various findings can be seized upon at will by those pundits who wish to "use" research evidence to "prove" their point.

One is, for example, somewhat irritated by the implications which are drawn for teacher trainers; to illustrate the sort of point that has to be made, the authors state that "at one extreme it could be argued that self-direction is the best way to facilitate learning, while at the other extreme it is often claimed that direct instruction is best" (p. 268). They go on to note that their data support neither of these extreme positions. That is, they resort to the device of setting up a straw man because it must be rare to find a teacher trainer who advocates either of these extreme positions. I suspect that the vast majority of teacher trainers attempt to introduce students to practices which are remarkably similar to those characteristically found in effective schools as described in *School Matters*. It is not beyond the bounds of possibility that teachers behave in these ways because they were taught to do so during their training programmes. Further on, they recommend that students should not be trained to organise too many different learning experiences within any one classroom at any one time. I wonder how much experience the authors have of teacher training, because it is not typical to witness a student on school practice organising a whole variety of different learning experiences during a particular lesson, and the reason why students don't do this is precisely because of the difficulties which are indicated by the researchers themselves.

## CONCLUSION

I would not wish to argue particularly with the central findings of this research — indeed, I tend to agree with many of them. I would accept that there is substantial evidence to indicate that especially as far as children from the more deprived inner city areas are concerned learning is likely to be facilitated if they are taught in secure, well-structured environments in which the head and the teachers have got clearly articulated and agreed goals. Children will learn better if they are in classrooms which are work-orientated and in which teachers carefully monitor their progress. But having said this, it will only be when specific hypotheses investigating effective schooling have been subjected to empirical tests under experimental conditions that researchers will be able to claim that they have superior knowledge, or knowledge based upon a more secure foundation than the informed layman.

## REFERENCES

BRAUN, C. (1987). Teachers' expectations. In DUNKIN, M. J. (Ed.), *The International Encyclopaedia of Teaching and Teacher Education*. Oxford: Pergamon Press.

BENNETT, N. (1976). *Teaching Styles and Pupil Progress*. London: Open Books.

BENNETT, W. J. (1987). *What Works: Research about Teaching and Learning* (2nd ed.), United States Department of Education.

BIDDLE, B. J., and DUNKIN, M. J. (1987). Effects of teaching. In DUNKIN, M. J. (Ed.), *The International Encyclopaedia of Teaching and Teacher Education*. Oxford: Pergamon Press.

C.I.P.F.A. (1984) *Education Statistics, 1984-1985 Estimates*.

COLVIN, R. (1987). Cultural comparisons of schooling. *Educ. Researcher*, 16 (6).

EDUCATION (1988). *News report*, Vol. 171, p. 311.

GLASS, G. V. (1987). What works? Politics and research. *Educ. Researcher*, 16 (3).

MORTIMORE, P., *et al.* (1988). *School Matters: The Junior Years*. London: Open Books.

RABINOWITZ, W., and TRANES, R. M. W. (1953). Problems of defining and assessing teacher effectiveness. *Educ. Theory*, 3, 212-219.

RUTTER, M., *et al.* (1979). *Fifteen Thousand Hours*. London: Open Books.

SHULMAN, L. S. (1986). Paradigms and research programs in the study of teaching: a contemporary perspective. In WITTROCK, M. C. (Ed.), *Handbook of Research on Training* (3rd ed.). London: Collier Macmillan.

WINEBURG, S. S. (1987). The self-fulfilment of the self-fulfilling prophesy. *Educ. Researcher*, 16 (9).

ZIRKEL, P. A., and GREENWOOD, S. C. (1987). Effective schools and effective principals: effective research? *Teachers College Record*, 89 (2).

DESFORGES, C., and COCKBURN, A. (1987). *Understanding the Mathematics Teacher.* Basingstoke: Falmer Press, pp. 167, n.p., pbk., ISBN 1-85000-213-4.

Concern about the apparent inadequacy of much of our mathematical thinking (as opposed to responding by routine) has fuelled this sensitive and readable study by Desforges and Cockburn. But these two educationists are mindful of all that has gone before, from the comments of James and of Dewey through to the findings of Nuffield and Cockroft. So, mercifully, the reader is not treated to naive explanations of failure, but to serious attempts to inquire and reflect. Furthermore, one picks up the message that these researchers are themselves experienced teachers and ones who display empathy with and awareness of real classroom life. Thus, when dealing with questions like "why don't teachers take good advice?", the writers do not start from a basis which assumes teachers to be wilfully unaware of the need that effective understanding be grounded in practical investigation and experiential learning. Lack of awareness is not the problem, they say. But developing that awareness into appropriate management and organisational responses, rather than holding it in insulation from the normative day-to-day system maintenance, certainly is.

The authors briefly review the problems inherent in classroom learning generally and, whilst concentrating on illustrations related to teaching mathematics, emphasise that "tasks with higher level cognitive demands" increase risk and alter the "established exchange rate" in classrooms. The actual field work, with seven experienced infant teachers and involving some 296 interviews with children, gives a picture of teachers that is consonant with the authors' earlier expressed views. Teachers *were* aware of difficulties, were concerned that the children experienced and internalised real applications, but were also acutely aware of their own limitations in managing child-child interaction and a high level experiential approach. Tellingly and disappointingly the teachers seemed to doubt the child's capacity for listening productively. However, individual differences were clearly never far from the teachers' minds and the constant challenge of catering for the different entering characteristics of each learner very much to the fore.

As "outsiders" the authors' observations are acute and compelling. Industry and application were evident in these infant classrooms, but practical work intended to be used in a "supportive role to pencil and paper work". The children themselves seemed much less interested in discussions.

Overall, the book makes unusually absorbing reading. For this reviewer it rates more highly than many educational text-books he has read in recent years. Why? Because this little volume unpacks the real issues with classic simplicity and immediacy. It "speaks to one's condition", as the Quaker saying runs. It shows how collusive are some of the factors which make classroom learning what it is. It shows how difficult it is to "deliver" a more interactive, enquiry-based maths curriculum for young children. Unlike the current vogue, it does not blame the teacher, but argues for placing the major responsibility where it belongs. LEAs, HMIs (and our own Shell Mathematical Centre?) please note.

PHILIP GAMMAGE

---

EISENBERG, N., and SAYER, J. (1988). *Empathy and its Development.* Cambridge: Cambridge University Press, pp. 406, £32·50 hbk., ISBN 0-521-32609-5.

Few people would disagree with the editors of this book when they argue that "among the many self-evident problems in modern society are the lack of communication and understanding between individuals and groups and the frequent occurrence of inhumane and uncaring behaviour". A prime reason for this lack, the editors imply, is that we insufficiently study or teach one of the vital factors that leads to communication and understanding, namely empathy. So far so good. But although most of us probably feel we have experienced empathy at least at some point in our lives, how on earth can we define the term in a form that allows for study, and ultimately for encouragement and development in ourselves and others? The editors confess the concept is "a slippery one", and then manage to come up with a short and rather good definition themselves, namely that empathy is (or rather involves) "sharing the perceived emotion of another — 'feeling with another'".

One of the strengths of this definition is that it stresses empathy involves sharing a *perceived* emotion. Our perception of what another person is experiencing at the affective level may or may not be correct, but the important thing is that we think we recognise the nature of this experience, and evoke in ourselves our own equivalent of it (sympathy differs from empathy, the editors suggest, in that it involves "feeling for" another instead of "feeling with"). The definition therefore immediately opens up experimental possibilities. It authorises us to ask subjects what they perceive the affective states of other people to be, and to enquire whether this perception evokes matching states in themselves. True, there is a heavy reliance here upon self-report, and obvious problems when it comes to assessing the relative quantity of affect between subjects, but it is a start at least. And it avoids the difficult question, initially at least, of whether or not subjects are *accurate* in their perceptions of what others are experiencing. You don't necessarily have to get it right in order to be empathic, you simply have to possess the capacity to respond to others by going through what they seem to be feeling.

Having set out their stall, the editors divide their book into four parts. The first covers the theoretical aspects of the concept of empathy and the historical background to its study, the second looks at early empirical work, the third at current research projects, and the fourth examines the strengths and weaknesses of the various methods of operationalising empathy. The first thing that will strike many readers is the sheer volume of work that has already been done. The overwhelming majority of it comes from North America (are we in Britain too reserved to admit publicly the existence of empathy?), and this could be a disadvantage, since the experience of empathy may be very much a culturally determined thing. But there are certainly enough ideas generated by the work to indicate its psychologically respectability, and the publication of this book could start off a new (and welcome) fashion in experimental investigation.

Amongst the many topics covered are empathy and moral judgment, empathy and psychotherapy, empathy in children, gender differences in empathy (men may not be that much less empathic than women), age difference (no consistent pattern after age 11), parent/child empathy, temperamental and familial issues, altruism, and motor mimicry. To get away from simple self-report, somatic and physiological indices are also covered, and there are perceptive critical overviews for each of the four sections of the book by such luminaries as Paul Mussen. Things are rounded off by a good concluding chapter from the editors themselves, in which they admit that research generates even more questions than are originally posed. Their hope that their book will help towards the clarification, integration and evaluation of the whole area of empathy study is, however, soundly based. I welcome their work without reservation. It represents an important step forward in establishing empathy as a legitimate area for scientific psychological investigation, and thus an important step forward in the humanising and liberalising of the discipline of psychology as a whole.

DAVID FONTANA

---

FRIEDMAN, S. L., SCHOLNICK, E. K., and COCKING, R. R. (Eds.) (1987). *Blueprints For Thinking: The Role of Planning In Cognitive Development.* New York: Cambridge University, pp. xv + 559. £40·00 hbk., ISBN 0-521-25605-4.

The more one thinks about the potential role of planning in the behaviour of man and animals, the less confident one becomes about what it is. In one sense any pattern of behaviours which are manifestly not random might be deemed to be "planned", but that begs too many questions.

The first two sections of this book deal in a variety of interesting ways with these questions. To what extent should we require that planning is a conscious activity? Is planning a uniquely human activity? Is planning prevalent throughout our day-to-day living, or is it reserved to denote our responses to serious problems? The discussions of these issues are lucid and thought-provoking, but the reader should not expect a definitive statement to emerge, for it does not.

Hence the frame of reference with which one embarks upon the rest of the volume is still loose, if not tending to the amorphous. It is not clear, for example, what is to be gained from describing some simple act such as screwing up a bottle cap (with its overtones of the seminal and frequently cited TOTE mechanism proposed by Miller *et al.*, 1960) as an exemplar of the same phenomenon as a complex fishing routine with interchangeable and non-interchangeable subroutines. And the issue is confounded further when one discovers that other writers in the volume are arguing that analyses should not adhere strictly to logical models, but should incorporate motivational, emotional, social, linguistic, and metaplanning variables too.

The chapters on the development and deployment of planning in childhood revisit some of the difficulties psychologists have faced elsewhere. It would seem that the behaviour of all school-age children gives clear indication of planning, although younger children have understandably fewer resources on which to plan than their elders and at least one contributor proposes that planning is fairly ineffectual before adolescence. However, many of these studies seem to give little attention to the possible effects of confounding variables, particularly language and knowledge.

This is an interesting approach to the psychology of problem solving (for that is what it is). It is, on the whole, well written, though at least one author makes the reader suffer what Andrew Ellis has cleverly called "the bins and arrows" of outrageous diagrammatic model making. It draws together a number of related approaches in an useful way, and offers much food for thought.

GEOFF BROWN

## REFERENCES

MILLER, G. A., GALLANTER, E., and PRIBRAM, K. H. (1960). *Plans And The Structure Of Behaviour.* New York: Holt.

RAPOPORT, R. N. (1987). *New Interventions for Children and Youth. Action-Research Approaches.* Cambridge, New York: Cambridge University Press, pp. 275, £27·50. ISBN 0-521-34122-1.

HURRELMANN, K., KAUFMANN, F., and LOSEL, F. (Eds.) (1987). *Social Invervention: Potential and Constraints.* Berlin, New York: Walter de Gruyter, pp. 399. DM.120, US$49·95. ISBN 0-89925-327-X (US), ISBN 3-11-011256-6.

Applied social research is about intervening in the lives of others. It differs from applied research in the so-called natural or "hard" sciences, in that the "materials" to be worked with are human. And human beings resist, object, judge, choose, grow angry — in short, intervention in the social as opposed to the "natural" realm is infinitely more complex. That is why any book which aspires towards enriching the debate on the theory of how to intervene is welcome. The two books I have been asked to review take the problem of applied research from a different tradition and background of experience from my own as an educationist, grounded in the phenomenological, action research traditions of educational research and the democratic evaluation perspectives espoused by CARE*. For that reason I found them doubly interesting. Because for reasons of space, I cannot do either book full justice here, I should like to select those aspects I found particularly useful.

Rapoport approaches social intervention through action research and provides a brief history of the term and a range of approaches in the first chapter. Action research is interdisciplinary by nature. Although it is only one of many ways of studying innovations, "properly conducted, the scientific gains through involvement in a new intervention program can have advantages which the more classic experimental research models lack, particularly in developing new paradigms for understanding social change processes" (p. 29).

Indeed, he considers, some situations are so new that only through action research can knowledge be gained about them. And this provides the theoretical theme to the book. The substantive theme is "children and youth" in relation to those institutions or contexts which intervene in some way in their lives: family, school, courts, and communities. He provides examples of the problems faced by action researchers who in attempting to analyse a particular intervention programme "anticipated the possibility that they would have negative results requiring them to criticise persons and agencies with whom they had developed intimate association" (p. 53).

---

* Centre for Applied Research in Education, University of East Anglia.

This is a problem characteristic of evaluation research. It is essentially the problem of audience, an audience having a personal, emotional commitment to the programme itself. How does one feed back research information to practitioners in ways which are constructive, which do not create resistance to change and which maintain the support of the practitioners? The relationship between action personnel and research staff is critical. For many practitioners research is a "dirty word": in such a situation, how does one make the relationship productive? By presenting accounts in each of the substantive areas of case studies emphasising the constructive role of feedback (providing formative evaluation) and of forms of collaboration Rapoport provides an insight into how such problems may be handled. Also, such an approach will enable the construction of theory grounded in practice. As Rapoport puts it, action researchers give up a modicum of purity in research design but gain "through engaging in the more turbulent world of action, a sense of drama and relevance in their work and the possibility of using their knowledge to benefit humanity" (p. 260).

Hurrelmann, Kaufmann and Losel's book is different, not simply in that it is an edited book, but in that its concerns are wider. The approaches of the contributors are varied in style and content. Kaufmann, in the introduction, seeks to offer a logic of social intervention whilst rejecting the traditional causal model of prevention and intervention. The idea is to understand the processes of intervention in order to provide "the active participants of this process with insights into the way in which measures initiated by them may become effective" (p. 12). The ideas are developed in the often heavy theoretical language of structuralism, a language which pervades much of the book.

It is a relief to read the almost "folksie" article by Albee discussing powerless and politics within the realm of social intervention. It makes no sense, he argues, for us to develop separate programmes for various exploited groups. What they all have in common is that they are made to be powerless by an exploitative society. Thus, "If we see all these groups as powerless because of the operation of exploitative socio-economic forces, then the logical approach would be to determine ways to achieve an equitable redistribution of economic power". The task for any intervention programme is to seek ways to empower people. He ends, "I urge Community Psychology to formulate a basic manifesto, a statement of principles, that will announce our alliance with the exploited and offer our assistance to all those who struggle for a redistribution of power" (p. 51). It is asking for a total system approach.

A broadly system-theoretic approach pervades the book, each chapter focusing upon a particular application or development of the approach. But it is the theme of empowerment which attracts me. It is taken up again by Cochran. He identifies the problem as being that society has embraced the advantages of the capitalist system without having also accepted responsibility for the disadvantages or "fall-out": "power discrepancies between management and labour, relatively high levels of unemployment . . ." (p. 105). Several chapters explore the theme that the structure itself is an agent in generating problems. But between structure and individual there is an interplay, where the individual is seen as powerful (having a hold of society's levers) or as powerless. Empowerment thus becomes a crucial concept for intervention theory.

While the themes discussed are important and the theories developed interesting, it seems to me that empowerment can only be properly approached through forms of research which observe the individual and the social group in action. Such theories should rather be grounded in everyday practice and experience rather than grounded in metaphors of system. It is here that I found the books disappointing. There is little attention to the real scenes of everyday social life. Analysis takes place at a macro and theoretical level. It is, of course, vital to have a theoretical conception of systems but equally it is important not to lose sight of the individual's life which is never simply either a relationship (in the Marxist view) nor simply an "element". Both books lack sufficient concrete portrayal of life in action. And it is individual lives that are the objects of social interventions. That said, I happily recommend both books to those genuinely interested in the study of social interventions.

JOHN SCHOSTAK

SIEGEL, H. (1988). *Educating Reason: Rationality, Critical Thinking and Education.* London: Routledge, pp. 191, £22·50 hbk., ISBN 0-415-00175-7.

This book presents a sustained argument for critical thinking as a fundamental educational ideal with wide ranging implications. Siegel claims that despite considerable interest in critical thinking, partly inspired by the "Informal Logic Movement", the notion remains obscure. He therefore starts with an analysis and review of three currently popular accounts of critical thinking provided by Robert Ennis, Richard Paul and John McPeck. There follows Siegel's own positive account (the "reasons conception"), according to which the critical thinker is one who is appropriately moved by reasons. This conception emphasises the connection between critical thinking and rationality (critical thinking is seen as the "educational cognate" of rationality) and also the importance of dispositions, habits of mind and character traits, in addition to reasoning skills, for a full conception of critical thinking.

Siegel goes on to explore and tackle the problems involved in justifying critical thinking as an educational ideal and to present and discuss two possible objections. According to the "ideology objection", critical thinking, like other educational aims and ideals, can only be justified within the context of some ideology: "neutral" justification is impossible. The "indoctrination objection" concerns the threat to the ideal which arises if it turns out that education for critical thinking can only be conducted in an indoctrinative way. Siegel argues, however, that both objections fail.

The final two chapters aim to demonstrate the power of the ideal in respect of its implications for educational policy and practice. Science education is taken as a representative curriculum area: Siegel argues that critical thinking conflicts with a Kuhnian-style education for "normal science", that the ideal provides a powerful basis for evaluating alternative conceptions of science education and that it supports a pluralist approach which aims to help science students to become "fluid enquirers". Chapter 7 examines minimum competency testing which, Siegel argues, is incompatible not only with critical thinking but with any defensible conception of educational aims. His postscript highlights unresolved problems and points to the need for a general theory of rationality to ground the analysis of critical thinking.

This is an informative and well organised contribution to the debate about rationality and education and it provides a welcome introduction to the American literature in the area. I have reservations about the merits of some of the arguments (for example, the evolutionary analogy on pp. 134-5, doubts about which are in fact acknowledged in a footnote, and the transcendental arguments which appear on pp. 132 and 137), but there is rarely any doubt about what Siegel is saying, since the book is written with exemplary clarity. I learned much from the discussion of different conceptions of critical thinking (chapter 1), a discussion which has, among other merits, that of warmly acknowledging the positive contribution to the debate made by other critical thinking theorists.

As Siegel's postscript indicates, many problems remain to be addressed: in fact the conception of critical thinking presented is formal (as opposed to substantive) because of the lack of an underlying theory of rationality. Siegel claims that the solution of various problems connected with rationality is independent of their ramifications for critical thinking but it is hard to believe that the tasks of, for example, justifying critical thinking as an educational ideal and drawing policy and curriculum implications will not partly depend on *how* the conception of critical thinking is fleshed out. The postscript emphasises the inadequacies of instrumentalist (means-ends) theories of rationality: this strongly suggests that the work of critical theorists (the "Frankfurt School") would be a useful resource in the further development of ideas about critical thinking. (Strangely, the critical theory literature is virtually absent from the bibliography, although Habermas is mentioned in a footnote.)

Siegel states that an implicit aim of his book is to demonstrate the need for critical thinking theorists and educational policy and curriculum makers to attend to underlying philosophical issues. In showing the relevance of philosophical questions concerning rationality to a whole range of educational issues, he succeeds admirably in achieving this aim.

CAROLYN STONE

TOPPING, K. (1987). *The Peer Tutoring Handbook*. Beckenham: Croom Helm, pp. 123, £16·95 hbk., ISBN 0-7099-4349-2; £6·95 pbk., ISBN 0-7099-4353-9.

As someone interested in the whole area of co-operative learning methods and their implementation in classrooms I opened this book with some interest. The blurb had heightened this by promising that the book brought peer tutoring into the realm of practice of every teacher, and that it incorporated a step by step guide to organising this, using a minimum of time, effort and special materials. However, that interest soon turned to confusion.

The first confusion concerned the difficulty in acquiring an understanding of what peer tutoring was; what the term did, and did not, encompass. This was never resolved since a definition was never offered, and, to exacerbate the problem, the focus swung from what is conventionally understood as peer tutoring to related areas of group work, paired reading, reciprocal teaching and the like.

The second confusion concerned possible explanations of the claimed effects. No self respecting teacher educator would urge teachers to implement an innovation which was theory-free; practice without theory is blind. Yet in this book tutoring is placed within no clear theoretical perspective, and the only explanation for its claimed effects is, to say the least, naive. The bald claim is that it works by "promoting increased time spent on task", irrespective, presumably, of the type, appropriateness and quality of the tasks involved.

The third confusion arose in relation to how this apparently ill-defined and poorly understood technique was to be implemented. The prescription provided is very behaviouristic. A mastery approach is suggested utilising highly structured, sequenced, step by step materials which would, at least, claims the author, save the teacher time and thought. But it also appears that the tutors have to be trained in such things as task presentation, skill demonstration, assessing pupil performance, providing feedback, identification of consistent error patterns, the development of alternative learning experiences — all teaching skills which recent research has shown to be both critical, yet often poorly developed among experienced teachers. Whilst accepting that these are sophisticated skills the author totally fails to appreciate the impossibility of carrying out such training effectively in the classroom setting. Nor is implementation helped by the planning guide provided in chapter 5. This is no simple step by step approach as promised in the blurb, but a list of well over one hundred things to consider when planning a peer tutoring project. Overall the problems of implementing innovation are either ignored or not taken sufficiently seriously.

The fourth confusion arises from the pattern of the book itself. Not only is the sequence rather strange, e.g., research on peer tutoring is left until the penultimate chapter, and the level sometimes inappropriate, particularly the chapter on evaluation which is totally unsuitable for a teacher audience, the book as a whole does less than justice to its topic. I happen to think there is some pedagogical mileage in peer tutoring arrangements but this book will help to put few people on that road.

NEVILLE BENNETT

WHITE, R. (1988). *Learning Science*. Oxford: Basil Blackwell, pp. 227, £8·95 pbk., ISBN 0-631-15699-2.

Richard White's valuable and timely book is addressed to educational psychologists and teachers, but it is not just another book about teaching. As its title indicates, it is concerned less with what teachers do than with what is going on inside pupils' heads when they seek to understand what they are taught in the science classroom and laboratory. His central thesis is that current teaching practice rests largely on unstated but historic and unexamined assumptions about the ways in which students learn and that, if teaching is to advance from a craft to a profession, those assumptions must be replaced by an explicit and coherent theory of learning which has predictive value.

Although the model presented is described as being new, it is, in essence, an information processing model and many of its components — attention, perception, the establishment of long term memory, chunking, etc., will be reassuringly familiar to educational psychologists. What is new is his use of the model to provide a critical analysis of pupils' learning experiences and needs.

However, he is not content merely to illustrate the inadequacies of many current instructional strategies; he shows how the abstractions of learning theory can be embodied in the realities of classroom activities.

In discussing the changes in teaching practice predicated by learning theory, White is unashamedly prescriptive. Nevertheless, teacher readers will be enlightened and reassured rather than admonished by his arguments. Drawing on his own experience as a teacher, and from research studies, White argues that students' learning *can* be improved, for example, by taking account of the limitations of short term memory, by improving the context of learning and by reducing the ambiguities of perception and communication. Above all, he stresses the purpose of science education, that pupils should *understand* and not merely *know* Accordingly, he suggests the use of procedures to probe understanding which are novel to the classroom, such as concept mapping and the use of Venn diagrams and diagnostic interviews. The examples in the text are relevant and familiar to science teachers, and the illustrative anecdotes are pertinent, witty and well presented.

In the present state of instructional practice, there are perhaps two criteria against which a book on the application of cognitive psychology in the classroom should be judged. The first is its usefulness in assisting teachers to question the "deficit model" of the pupil, according to which pupil learning failure is interpreted as the result of deficiencies in the pupil, e.g., of intelligence, application, or attainment of the presumed necessary level of reasoning. Such a model is stultifying since it leads to a fatalistic rather than an interventionist attitude among teachers. Unfortunately, it seems rather too easy for anyone reading *Learning Science* without the care it deserves to persist with the deficit model (since, as White notes, "what you know determines what you see"), merely by replacing the traditional deficits with "lack of intellectual skills' such as chunking, discriminating, classifying, etc. It is only by paying close attention to White's argument that the reader will be made aware that these skills are critically dependent on the richness of the student's knowledge structures and history of successful experience within the domain, both of which are significantly dependent upon the wisdom and skill of the teachers.

The second criteria for a book in this genre is its usefulness in motivating educational psychologists to reduce their preoccupation with child guidance, dealing with the problems of the few children whose education has significantly gone off course, and to address themselves to the improvement in learning of the many children in every classroom whose education is proceeding "normally" — and badly.

It may be thought that the task of developing the expression of cognitive psychology in the classroom is not merely straightforward but one for teachers alone. In fact, the changes in aims and practices which White suggests are the fusion products of his theoretical insight and of his practical experience as a teacher. The necessity for psychologists to work along with teachers in order to develop a psychology of the school subject is implicit, but the clarion call is missing.

These strictures aside, this is a useful and thoughtful book, well written and well produced, which deserves a wide readership among all educationalists.

MARY SIMPSON

---

WILLIAMS, P. (Ed.) (1988). *A Glossary of Special Education.* Milton Keynes: Open University Press, pp. 211, £8·95 pbk., ISBN 0335-15995-8.

Over the last ten years professionals have increasingly responded to the need to work with colleagues and parents in meeting the special needs of some children. Effective communication has been essential both in establishing a multidisciplinary approach and the understanding as to how some children's needs may be both the same and different from those of all children. It could be argued that the use of specific terminology has been a barrier to that effective communication. A glossary of such technical terms would seem a laudable attempt to break down such a barrier.

The glossary is intended, therefore, as a response to the Warnock Report and the 1981 Education Act by developing a database of multidisciplinary terminology currently used by professionals working in the field of special education for use by parents, professionals, managers and administrators.

The book is the result of a two-year project, begun in 1983, supported by the Manpower Services Commission's Community Programme. Terms were originally culled from the index of the education library at the University College of North Wales. The criteria for inclusion were based on the frequency of appearance in the original screening process, its current usage, its relevance for a wide audience and the expectation of its international usage. Some terms were retained if the interest was limited but if the term was highly significant to that group.

A further two years of the project enabled the writers to revise the databases, define the terms against the sources in which they appeared as well as testing out responses from individuals in order to monitor the level of detail, difficulty and accessibility. The glossary does include earlier legislation for historical interest and testing procedures which were once widely used.

In general, the definitions are in a clear readable style with commendable care taken not to imply value judgments in this often complex field of inquiry. The cross-referencing, with a few minor errors, does allow the reader to share the search for information in an enjoyable way.

Searching for understanding, however, depends on what you bring to the situation. As this book is primarily aimed at parents, the effectiveness of the glossary would seem to be for the definition to inform with clarity and for the cross-referencing to lead the reader to a deeper and wider understanding. Yet parents of children with special needs do not, one could argue, read such terminology in an objective way but are highly sensitive to and interpretive of language that, even for informed professionals, operates at a subjective level.

The term maladjustment, for example, provides the reader with a definition, knowledge of the Underwood Report, exclusion from the 1981 Act and the preferred term e.b.d. Cross-referencing leads to neurotic and anti-social behaviour, delinquency, special class, child guidance, schools' psychological service and psychotherapy.

With the aid of the glossary a parent might well be better informed about the term maladjustment but confused in relation to their own child. Such confusion could provide a platform for dialogue but this would be better dealt with in a supportive situation. This glossary would, therefore, be an invaluable aid to voluntary groups, societies and trusts who aim to support children with special needs. It would be devalued if used as a "home doctor", especially in view of its attention to medical conditions and testing procedures. Such a usage would be contrary to the purpose of the glossary and the developing notion of "special needs".

Indeed, one could argue that a dialogue about "needs" eliminates the need for technical language other than that related to provision for the meeting of those needs. It is here that one sees omissions in the glossary. Parents and paraprofessionals are much more likely to encounter terms such as match, profiling and alternative curriculum. This weakness may have arisen from the glossary's original construction. The editor, however, welcomes readers' views on omissions, difficulties and principles of usage and, presumably, updating.

The glossary is a useful source of legislative, medical and educational terminology which, at its best will serve to clarify and illuminate but which could, with less informed usage, serve to maintain and promote the identification of differences which demand specialist responses and special provision.

MARILYN TUCKNOTT

ZARKOWSKA, E., and CLEMENTS, J. (1987). *Problem Behaviour in People with Severe Learning Disabilities.* Beckenham: Croom Helm, pp. 208, £10·95 pbk., ISBN 0-7099-3014-3.

This text is intended to be a practical manual for care staff (residential, teachers, psychologists, nurses) who are involved in the long term resolution of behaviour problems in people with severe learning disabilities. Its intention is "to help staff understand the meaning of problem behaviours for the individual, and on the basis of this to develop broad based interventions using a combination of well established behavioural techniques and innovative ecological approaches". The chapters lead the reader through the basic rationale to a system for establishing a management plan. This is followed by a discussion of problem behaviour reduction and the teaching of alternative skills. Later chapters focus on two key issues of self-management and high quality practice. The text ends with useful reference material relating to case study examples, recording charts and assessment schedules.

The authors are clearly experienced in dealing with a variety of problem behaviours arising from children and adults with severe learning disabilities. They describe their approach as both individualised and multi-dimensional. It is firmly rooted in the behavioural approach, albeit in a more enlightened and up-to-date form. Their approach, therefore, has the advantage of a system which deals with observable behaviour in a consistent and structured manner, often valuable with problems which arise in this "client group". It does, however, provide the authors with problems in examining the cognitive and emotional factors which they clearly wish to include in their model. Their list of "strengths and deficits" uses the alternative label for the latter of "needs". Nowhere does there seem to be a recognition of this paradox.

The emphasis throughout is on a more enlightened behavioural approach so one is not surprised to read about "more normalised" environments and of a greater emphasis being placed on antecedent intervention. Each chapter outlines the general principles involved before giving practical step-by-step advice which is directly useful to the practitioner. The basic model used throughout is that of STAR, namely, settings, triggers, actions and results.

The authors advise a graded approach to intervention strategies, with consistent monitoring and evaluation in the context of a team approach. There is support for the rigour of the behavioural approach in the endorsement of the terms consistency, precision and objectivity. It is therefore surprising to read various comments recorded in the examples of check lists and programmes such as, "Has a nice smile, sings well, behaves nicely when on his bicycle, tell her she's been naughty, Philip stop that awful wailing". It does suggest a gap between what is advocated and actual practice.

Another criticism relates to the advocacy of over-correction (which combines both restitution and repeated positive practice) which I consider to be both unethical and counter-productive as a procedure. The section on the issues of "constraint" is too brief in view of its importance as a topic.

The chapter on self-management is helpful although the fact that error-free learning may inhibit the very skills which one is trying to enhance goes unremarked. This is followed by a detailed practical guide to recording progress by charts, check lists and graphs, which is helpful and clearly presented.

The final chapter concerned with "high quality practice" is valuable and deals with such issues as ownership of the problem, team approach, leadership and system evaluation.

Is the text worth purchasing? The answer must be in the affirmative but with the qualification that it is predominantly behavioural in approach and intended to be of greater use to specialist professionals in clinical and educational settings.

PETER W. GURNEY

# PUBLICATIONS RECEIVED

Books for review have been selected from the following
("n.p." means "no price mentioned)

Bernhard, J. G. (1988). *Primates in the Classroom.* Amherst, Mass.: University of Massachusetts Press, pp. 204, £22·00 hbk., ISBN 0-87023-610-5.

Black, H. D., and Dockrell, W. B. (1988). *New Developments in Educational Assessment, Monograph Series No. 3.* Edinburgh: Scottish Academic Press, pp. 186, n.p.

Butler, D. (1988). *Babies Need Books.* London: Penguin, pp. 212, £3·95 pbk., ISBN 0-14-010094-6.

Byrne, E., Cunningham, C., and Sloper, P. (1988). *Families and their Children with Down's Syndrome.* London: Routledge, pp. 159, £8·95 pbk., ISBN 0-415-00608-2.

Cormier, S. M., and Hagman, J. D. (1987). *Transfer of Learning.* San Diego: Academic Press, pp. 281, $39·50 hbk., ISBN 0-12-188950-5.

Crystal, D. (1988). *The English Language.* London: Pelican, pp. 288, £4·95 pbk., ISBN 0-14-022730-X.

Egan, K. (1988). *Teaching and Storytelling.* London: Routledge, pp. 122, £6·95 pbk., ISBN 0-415-00700-3.

Eisenberg, N., and Glasgow, D. (1988). *Issues in Clinical Psychology 1986.* Aldershot: Gower, pp. 218, £25·00 hbk., ISBN 0-566-05635-6.

Furneaux, B. (1988). *Special Parents.* Milton Keynes: Open University Press, pp. 103, n.p. ISBN 0-335-15122-1.

Gupta, R. M., and Coxhead, P. (1988). *Cultural Diversity and Learning Efficiency: Recent Developments in Assessment.* Basingstoke: Macmillan, pp. 192, £29·50 hbk., ISBN 0-333-41337-7.

Franklin, M., and Barten, S. S. (1988). *Child Language.* Oxford: Oxford University Press, pp. 398, £22·50 hbk., ISBN 0-19-504332-4.

Hull, R. (1988). *Behind the Poem.* London: Routledge, pp. 265, £10·95 pbk., ISBN 0-415-00701-1.

King, N., Hamilton, D., and Ollendick, T. H. (1988). *Children's Phobias.* Chichester: John Wiley, pp. 260, £27·50 hbk., ISBN 0-471-1027-8.

Krathwohl, D. (1988). *How to Prepare a Research Proposal.* New York: Syracuse University Press, pp. 302, $14·95 pbk., ISBN 0-8156-8112-7.

Kuhn, D., *et al.* (1988). *The Development of Scientific Thinking.* San Diego: Academic Press, pp. 249, $29·95 hbk., ISBN 0-12-428430-2.

Levenstein, P. (1988). *Messages from Home.* Ohio: Ohio State University Press, pp. 247, $22·50 hbk., ISBN 0-8142-0447-3.

Lyons, J. (1987). *Ecology of the Body.* London: Duke University Press, pp. 338, £35·65 hbk., ISBN 0-8223-0710-3.

McClelland, D. C. (1988). *Human Motivation.* Cambridge: Cambridge University Press, pp. 663, £19·50 pbk., ISBN 0-521-36951-7.

McNaughton, S. (1987). *Being Skilled.* London: Methuen, pp. 280, n.p. ISBN 0-416-01622-7.

Ouspensky, P. D. (1988). *Conscience.* London: Arkana, pp. 159, £4·95 pbk., ISBN 0-185063-102-6.

Robinson, G., and Maines, B. (1988). *They Can Because . . . Workshop Perspectives 3.* Birmingham: Association of Workers for Maladjusted Children, pp. 24, £2·00 pbk.

Rutter, M., Tuma, A. H., and Lann, I. S. (1988). *Assessment and Diagnosis in Child Psychopathology.* London: David Fulton, pp. 477, £25·00 hbk., ISBN 1-85346-058-3.

Salmon, P. (1988). *Psychology for Teachers.* London: Hutchinson, pp. 138, £6·95 pbk., ISBN 0-09-172950-5.

Schrag, F. (1988). *Thinking in School and Society.* London: Routledge, pp. 189, £22·50 hbk., ISBN 0-415-00174-9.

Spencer, K. (1988). *The Psychology of Educational Technology and Instructional Media.* London: Routledge, pp. 198, £22·50 hbk., ISBN 0-415-00567-1.

Sternberg, R. J. (1988). *The Psychologist's Companion.* Cambridge: Cambridge University Press, pp. 208, £7·50 pbk., ISBN 0-521-34921-4.

Sternberg, R., and Smith, E. (1988). *The Psychology of Human Thought.* Cambridge: Cambridge University Press, pp. 480, £15·00 pbk., ISBN 0-521-31115-2.

Weeks, D. J., and Ward, K. (1988). *Eccentrics: The Scientific Investigation.* Stirling: Stirling University Press, pp. 259, £16·95. ISBN 0-948812-00-1.

Weinstein, C., Goetz, E., and Alexander, P. (1988). *Learning and Study Strategies.* San Diego: Academic Press, pp. 353, $44·50 hbk., ISBN 0-12-742460-1.

White, R. (1988). *Learning Science.* Oxford: Basil Blackwell, pp. 227, £8·95 pbk., ISBN 0-631-15699-2.

# AUTHOR INDEX TO VOLUME 58, 1988

Brooks, M. R.: Pre-School Children's Preferences for Drawings of a Similar Complexity to their Own 165

Brown, R.: see Constable, H.

Brown, R.: see Speth, C.

Butler, R.: Enhancing and Undermining Intrinsic Motivation: the Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance 1

Campbell, R. M.: see Constable, H.

Casey, W.: Integration of Down's Syndrome Children in the Primary School: a Longitudinal Study of Cognitive Development and Academic Attainments 279

Clifford, M. M.: Failure Tolerance and Academic Risk-Taking in Ten- to Twelve-Year-Old Students 15

Clifford, M. M.: see Kim, A.

Coltheart, V.: see Laxon, V. J.

Constable, H.: Sectional Drawings from Science Textbooks: an Experimental Investigation into Pupils' Understanding 89

Cook, J.: An Investigation of the Validity of the British Ability Scales with Respect to the Weschler Intelligence Scale for Children — Revised and the Wide Range Achievement Test — Revised on a Group of Canadian Children 212

Corkill, A. J.: see Glover, J. A.

Craske, M. L.: Learned Helplessness, Self-Worth Motivation and Attribution Retraining for Primary School Children 152

Crozier, W. R.: see Brooks, M. R.

Cumming, G.: see Willich, Y.

Darom, E.: Sex Differences in Attitudes Toward School: Student Self-Reports and Teacher Perceptions 350

Dickson, G. C. A.: Children's Perception of and Anticipated Responses to Risk 147

Dobbins, D. A.: Yule's "Hump" Revisited 338

Entwistle, N. J.: Approaches to Studying and Levels of Processing in University Students 258

Foon, A. F.: The Relationship Between School Type and Adolescent Self Esteem, Attribution Styles, and Affiliation Needs: Implications for Educational Outcome 44

Francis, J.: see Green, L. F.

Glenn, S. M.: see Brooks, M. R.

Glover, J. A.: Detecting False Statements in Text: The Role of Outlines and Inserted Headings 301

Glover, J. A.: Remembering Written Instructions: Tab A Goes into Slot C, or Does It? 191

Gordon, C. J.: see Kirby, J. R.

Green, L. F.: Children's Learning Skills at the Infant and Junior Stages: a Follow-on Study 120

Harvey, A. L.: see Glover, J. A. (second entry).

Hattie, J.: Preferred Classroom Environment and Approach to Learning 345

Hewison, J.: The Long-Term Effectiveness of Parental Involvement in Reading: a Follow-up to the Haringey Reading Project 184

Hoffman, M.: see Schwarzwald, J.

Hunting, R. P.: Pre-Schoolers' Cognitions of Fractional Units 172

Hutchison, G. E.: see Dickson, G. C. A.

Jones, D.: see Casey, W.

Keating, C.: see Laxon, V. J.

Kim, A.: Goal Source, Goal Difficulty, and Individual Difference Variables as Predictors of Responses to Failure 28

Kirby, J. R.: Text Segmenting and Comprehension: Effects of Reading and Information Processing Abilities 287

Krug, D.: see Glover, J. A.

KUGLER, B.: see Casey, W.

LAXON, V. J.: Children Find Friendly Words Friendly too: Words with many Orthographic Neighbours are Easier to Read and Spell .......... 103

LEACH, C.: see Lister, C. M.

LISTER, C. M.: Sequence in Hearing Impaired Children's Development of Concepts .......... 127

McGEE, R.: Slow Starters and Long-term Backward Readers: a Replication and Extension .......... 330

McKAY, M. F.: Patterns of Performance on the British Ability Scales for a Group of Children with Severe Reading Difficulties .......... 217

McNAMARA, D.: Essay Review of *School Matters* by Mortimore *et al* .......... 356

MELDRUM, C.: see Worrall, N.

NEALE, M. D.: see McKay, M. F.

OWEN, D. R.: see Teasdale, T. W.

PRAMLING, I.: Developing Children's Thinking About Their Own Learning .......... 266

PRIOR, M.: see Willich, Y.

RICH, Y.: see Darom, E.

ROTEM, T.: see Schwarzwald, J.

SCHWARZWALD, J.: A Preparatory Programme for Passage into the Ethnically Integrated School: Impact on School Climate and Social Relations .......... 55

SHARPLEY, C. F.: see Hunting, R. P.

SILVA, D. A.: see McGee, R.

SØRENSEN, T. I A.: see Teasdale, T. W.

SPANOS, T.: see Willich, Y.

SPETH, C.: Study Approaches, Processes and Strategies: Are Three Perspectives Better than One? 247

TEASDALE, T. W.: Regional Differences in Intelligence and Educational Level in Denmark .......... 307

WATERSTON, S.: see Entwistle, N. J.

WATKINS, B.: see Casey, W.

WATKINS, D.: see Hattie, J.

WESENKRAFT, K.: see Lister, C. M.

WHITTINGTON, J.: Large Verbal — Non-Verbal Ability Differences and Underachievement .......... 205

WILLIAMS, S.: see McGee, R.

WILLICH, Y.: Are Disabled Readers Delayed or Different? An Approach Using an Objective Miscue Analysis .......... 315

WORRALL, C.: see Worrall, N.

WORRALL, N.: Children's Reciprocations of Teacher Evaluations .......... 78

ZEIDNER, M.: The Relative Severity of Common Classroom Management Strategies: the Student's Perspective .......... 69

# EDITORIAL ACKNOWLEDGMENT

The Editor and Assistant Editors of *The British Journal of Educational Psychology* gratefully acknowledge the assistance of the following, who have refereed papers submitted between 1st April 1987 and 31st March 1988:

M. Apter

N. Austin

N. Bennett

M. Beveridge

G. Brown

R. Burden

L. J. Chapman

A. Cline

R. Cowan

C. Cunningham

W. Curr

C. Elliot

N. J. Entwistle

P. Ernest

P. T. Farrell

T. Halil

D. Hamblin

P. Hannon

J. Harris

R. Harrison

M. Howe

M. Hughes

C. Hulme

P. Kline

C. Kyriacou

F. J. Macleod

A. Packer

A. Pinion

H. Pont

Linda Pring

C. Rogers

D. Rowntree

I. St. James-Roberts

L. Smith

Patricia Soper

J. Versey

N. Worrall

D. Wray

# NOTES TO CONTRIBUTORS

Manuscripts are accepted on the understanding that they report unpublished work that is not under consideration for publication elsewhere, and that if accepted for the *Journal,* the work will not be published again in the same form, in any language, without the consent of the Editor.

The usual range of length of papers is from 2,500 to 4,000 words. Papers should be in a clear style and free from jargon. Statistical matter (formulae, tables, discussion of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as mean, percentiles, standard deviations, correlation, chi-squared, standard error, critical ratio, significance and reliability, regression and factor analysis, can be employed in the text and need not be explained.

Short research notes will also be published, reporting work of two kinds: (i) results of sufficient moment to merit publication in advance of a more comprehensive paper; (ii) work which substantially confirms or extends existing knowledge but which does not justify an extensive paper.

Contributors can help to avoid delays in publication by observing certain guiding principles:

1—Papers should be written with the utmost conciseness consistent with clarity. If the editor or a referee takes the wrong meaning from or fails to understand a passage, the fault is the author's. Unless the editor can grasp the meaning of a sentence unequivocally, it cannot be assumed that other readers can

2—A paper should be written only when a piece of work is rounded off. A report of a pilot study, or a series of papers on the same subject written as the results come to hand, is unlikely to be accepted for publication. A comprehensive paper, particularly if it includes reports of several experiments or analyses, or replications on further samples, which give consistent results, will be given favourable consideration.

3—Experimental reports should generally follow the standard pattern of presentation, e.g..

(*a*) A short summary at the start.
(*b*) A statement of aims or objectives and a review of relevant previous work as an introduction to the topic—conciseness and strict selection of relevant aspects are particularly important here.
(*c*) A clear statement of the sample, the procedure, the tests, the experimental material, in sufficient detail to permit a reader to replicate the study if he so wishes.
(*d*) A factual statement of results, without speculative comment or discussion.
(*e*) Either a brief list of general conclusions, not repeating results already listed, or a discussion, which should be strictly limited to discussing the results (if this is necessary) and which should not involve mere recapitulation or speculation beyond the scope of the inquiry.
(*f*) Acknowledgments
(*g*) References.

4—The summary (which is often copied by abstracting agencies) should be intelligible without reference to the paper It should give an indication of the aim, scope and results of the inquiry, preferably in about three sentences. It should mention the size and nature of the sample, and should not include vague phrases such as "the results were analysed . . ." or "The implications of the findings are discussed ". Even for a full-length paper, the summary should not exceed 200 words.

5—Figures and tables should be selected to illustrate or summarise points made in the text The same data should not be reported both in table and figure, tables are generally preferable because they are much more economical. Tables should contain selected data: only the most important quantitative results can be reported in full.

6—Authors should consult a current issue of the *Journal* in order to follow the conventions applied in the *Journal*—the use of headings, numbering of tables, the absence of footnotes, the impersonal style of reporting, omission of dots in standard abbreviations, reference to subjects as children and not as Ss, the lay-out of references and so on. The British Psychological Society pamphlet, *Suggestions to Contributors,* should be consulted. References should be cited by author and date in the text, and arranged in alphabetical order in the list of references at the end All measurements should be stated in SI units The *Journal* is committed to the use of non-sexist language

7—A manuscript should be headed with the title of the paper, the names of the authors and, in brackets, a short address identifying the institution where the work was done Titles should be short If the title exceeds 45 letters and spaces in length, a shortened version, suitable as a running title, should be added.

8—The manuscript should be typed in double-spacing with wide margins on one side of the paper which is suitable for ink corrections and printer's instructions—not duplicated or on flimsy paper Drawings should be in Indian ink on cartridge paper Manuscripts which require re-typing before they can be handled by the printer have to be sent back to the authors. The *Journal* has no paid staff

9—Authors, if they specifically request, may have their papers reviewed blind

Authors receive 50 copies of their papers free (delivered about six weeks after publication). Extra copies may be had at cost price if the order is given when the proof is returned.

---

The *Journal* will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Authors who refer to physical measurements in their papers should now use SI units except where it would be absurd to do so. Further information about SI units, including conversion tables, is contained in the revised *Suggestions to Contributors* pamphlet issued by the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

## PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

*British Journal of Psychology*
*British Journal of Medical Psychology*
*Journal of Occupational Psychology*
*British Journal of Social Psychology*
*British Journal of Clinical Psychology*
*British Journal of Developmental Psychology*
*British Journal of Mathematical and Statistical Psychology*
*The British Journal of Educational Psychology*

The *British Journal of Psychology*, the *British Journal of Medical Psychology, Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1988 subscription prices, payable in advance, are £74·50 (US$134·00) for the *British Journal of Psychology;* £49·50 (US$89·00) for the *British Journal of Social Psychology*, £49·50 (US$89·00) for the *British Journal of Clinical Psychology*, £45·00 (US$81·00) for the *British Journal of Developmental Psychology;* £59·00 (US$106·00) for the *British Journal of Medical Psychology;* and £49·50 (US$89·00) for the *Journal of Occupational Psychology*

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £49·50 per volume post free (US$89·00 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts., SG6 1HN, UK. A special price of £340·00 (US$£610·00) for all seven journals going to the same address is offered for 1988.

*The British Journal of Educational Psychology* is issued three times a year, in February, June and November. The prepaid subscription is £20·00 UK, £22·00 overseas a year, post free, or £7·00 each number. The prepaid subscription price in the USA is $45·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

### MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D. E. Blackman, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Medical Psychology* should be sent to Dr. Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.

For the *Journal of Occupational Psychology* should be sent to Dr. D. Guest, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Social Psychology* should be sent to Dr. A. S. R. Manstead, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Clinical Psychology* should be sent to Professor R. Hodgson, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr. A. D. Lovie, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Developmental Psychology* should be sent to Professor George Butterworth, Journals Office, 13A Church Lane, London, W2 8DX.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Hazel Francis, University of London Institute of Education, 24-27 Woburn Square, London, WC1H 0AA.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

**All manuscripts should be submitted in duplicate.**

### OFFPRINTS

Contributors on application receive fifty copies of their papers free Additional copies may be obtained in units of fifty copies.

### REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., PO Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

### SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £40·00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £7·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOG

VOL. 58, PART 3, 1988

24 MAY

## CONTENTS

ARTICLES

SPETH, C., and BROWN, R. Study Approaches, Processes and Strategies: Are Three Perspectives Better than One?

ENTWISTLE, N. J., and WATERSTON, S. Approaches to Studying and Levels of Processing in University Students ........ 258

PRAMLING, I. Developing Children's Thinking About Their Own Learning ........ 26

CASEY, W., JONES, D., KUGLER, B., and WATKINS, B. Integration of Down's Syndrome Children in the Primary School: A Longitudinal Study of Cognitive Development and Academic Attainments ........ 279

KIRBY, J. R., and GORDON, C. J. Text Segmenting and Comprehension: Effects of Reading and Information Processing Abilities ........ 2

GLOVER, J. A., and KRUG, D. Detecting False Statements in Text: The Role of Outlines and Inserted Headings ........ 301

TEASDALE, T. W., OWEN, D. R., and SØRENSEN, T. I. A. Regional Differences in Intelligence and Educational Level in Denmark ........ 30

WILLICH, Y., PRIOR, M., CUMMING, G., and SPANOS, T. Are Disabled Readers Delayed or Different? An Approach Using an Objective Miscue Analysis ........ 3

McGEE, R., WILLIAMS, S., and SILVA, P. A. Slow Starters and Long-Term Backward Readers: A Replication and Extension ........ 330

RESEARCH NOTES:

DOBBINS, D. A. Yule's "Hump" Revisited ........ 338

HATTIE, J., and WATKINS, D. Preferred Classroom Environment and Approach to Learning ........ 345

DAROM, E., and RICH, Y. Sex Differences in Attitudes Toward School: Student Self-Reports and Teacher Perceptions ........ 3

ESSAY REVIEW:

McNAMARA, D. "School Matters" by P. Mortimore *et al.* ........ 356

BOOK REVIEWS ........ 36

PUBLICATIONS RECEIVED ........

AUTHOR INDEX FOR VOLUME 58 ........

EDITORIAL ACKNOWLEDGMENT ........

NOTES TO CONTRIBUTORS ........

UK ISSN 0007-0998